

Design and Synthesis of Covalent Kinase Inhibitors Utilising Novel Electrophiles

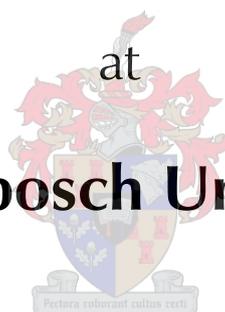
by

Luke Eric Hodson

*Submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy*

at

Stellenbosch University



Supervisor: Prof. W. A. L. van Otterlo

Co-supervisor: Dr. S. C. Pelly

Department of Chemistry and Polymer Science

Faculty of Science

December 2019

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly stated otherwise) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Luke Eric Hodson

Name in full

December / 2019

Date

Copyright © 2019 Stellenbosch University

All rights reserved

Abstract

Cancer is the collective name given to an extensive group of diseases which is mainly characterised by the uncontrolled growth of abnormal cells that either ignore or evade biological processes or display enhanced biological processes. Cancer is the second leading cause of death worldwide, claiming an estimated 9 million lives and accounting for 1 in 6 deaths in 2016. In a post-genomic era, rapid advances in molecular cancer biology and genomics has led to the identification of defining characteristics and traits of this complex disease, known as the hallmarks of cancer. Consequently, the last two decades has seen a shift from the use of conventional chemotherapy to the era of targeted cancer therapy.

Epidermal growth factor receptor (EGFR) is a transmembrane receptor which belongs to the ErbB/HER family of protein-tyrosine kinases. Owing to its expansive role in signal transduction pathways, EGFR regulates a host of essential aspects of cellular life, including division, growth and apoptosis. Oncogenic activation of this receptor, resulting in overexpression or hyperactivation, has been implicated in a variety of malignancies. To overcome this, researchers make use of tyrosine kinase inhibitors (TKIs) to effectively inhibit the receptor and cease uncontrolled cellular growth.

The past 16 years has seen the development of four generations of EGFR inhibitors attempting to overcome the emergence of multiple resistance mechanisms which are brought about by point mutations within the enzyme. Most recently, the emergence of resistance mediated by the EGFR-C797S mutation in 2015 has presented a challenging therapeutic endeavour. To date, no major breakthroughs have been achieved in targeting this clinically relevant mutant variant. As such, the research efforts documented within this thesis focus on the development of new therapeutic strategies to overcome this resistance mutation.

The first of these approaches entails the synthesis of highly potent reversible and irreversible EGFR inhibitors. Building on a previously identified lead compound, and drawing inspiration from dual EGFR/HER2 inhibitors, we incorporated several glycol chain functionalities at strategic positions to introduce favourable reversible interactions within the active site of the enzyme. Furthermore, studies on the effects of derivatisation and interchange of other structural elements, such as the solubilising group and acryl amide electrophile, were undertaken.

Departing from traditional inhibition methods, our second strategy sought the re-establishment of irreversible inhibition by targeting the EGFR catalytic Lys745 residue with novel electrophiles. To induce covalent bond formation, inhibitors containing 1,4-dicarbonyl warheads, able to undergo Paal-Knorr pyrrole formation with the Lys745 residue, were synthesised. Furthermore, potential reversible and irreversible interactions with the catalytic lysine residue were investigated with inhibitors containing nitrile and reactive sulfonyl fluoride moieties.

Lastly, we pursued covalent bond formation of the mutated Ser797 residue - the main mechanism by which resistance is conferred in the EGFR-C797S mutant variant. To accomplish this, inhibitors containing a multitude of reactive serine-targeting electrophiles were synthesised, including nitrogen-, boron-, phosphorous-, sulphur- and fluorine-containing warheads.

Biochemical and cellular evaluation of the 48 final compounds synthesised during this research project provided encouraging results. This included picomolar IC_{50} values for biochemical inhibition of the clinically relevant EGFR-L858R/T790M/C797S triple mutant variant of the enzyme by the synthesised reversible inhibitors. However, these results were unfortunately not reflected in the cellular-based assays. Additionally, several inhibitors containing the conventional acrylamide electrophile exhibited low nanomolar and picomolar range activity against the EGFR-L858R/T790M double mutant in a biochemical and cellular setting respectively, alluding to irreversible inhibition of the enzyme. A crystal structure of the most efficacious irreversible inhibitor complexed with the cSrc-T338M/S345C surrogate was obtained, providing confirmation of covalent bond formation with the Cys797 residue and reinforcing our postulates regarding the observed trends in activity. Finally, several lead compounds were appropriately evaluated and shown to be effective dual EGFR/HER2 inhibitors.

The use of covalent mass spectrometry experiments revealed a reversible binding mode for all inhibitors targeting the catalytic Lys745 and mutant Ser797 residue. However, several functional groups were identified for their ability to create favourable interactions with these residues, as reflected in their effective biochemical and cellular activity profiles against the EGFR-C797S mutant. This included the 1,4-dialkene-, boronic ester-, vinyl sulfonyl fluoride-, imidazole- and nitrile-bearing compounds.

Ultimately, novel synthetic methodology was established en route towards these inhibitors and their respective electrophiles. Highlights include development of a new and convenient one-pot installation of an acetamide- or propanamide-imidazole functional group to aniline derivatives utilising 1,1'-carbonyldiimidazole and the corresponding α -, β -halo-carboxylic acid. Furthermore, optimised methods were derived for future access to the variety of electrophiles and heterocyclic scaffolds synthesised throughout this research project.

Opsomming

Kanker is die kollektiewe naam wat gegee word aan 'n uitgebreide groep siektes, wat hoofsaaklik gekenmerk word deur die ongebeerde groei van abnormale selle. Hierdie selle vermy, of bevat negatiewe of verbeterde biologiese prosesse. Kanker is die wêreld se tweede grootste oorsaak van dood, wat ongeveer 9 miljoen lewens geeis het in 2016. Dit reken vir 1 uit 6 van die totale jaarlikse sterftes wêreldwyd. In 'n post-genomiese era het vinnige vooruitgang in molekulêre kankerbiologie en genomika gelei tot die identifisering van kenmerkende vermoëns van hierdie komplekse siekte. Gevolglik het die afgelope twee dekades 'n verskuiwing vanaf konvensionele chemoterapie na die era van geteikende kankerterapie gesien.

Epidermale groeifaktor reseptor (EGFR) is 'n transmembrane reseptor wat behoort aan die ErbB/HER familie van proteïen-tyrosien kinases. As gevolg van sy uitgebreide rol in seintransduksiepaaie, reguleer EGFR verskeie noodsaaklike aspekte van sellulêre lewe, insluitend verdeling, groei en apoptose. Onkogene aktivering van hierdie reseptor, wat tot ooruitdrukking of hiperaktivering lei, is in 'n verskeidenheid maligniteite betrokke. Navorsers gebruik tyrosienkinase-inhibeerders (TKI) om die reseptor effektief te inhibeer en ongebeerde sellulêre groei te staak.

Deur die afgelope 16 jaar is vier generasies EGFR-inhibeerders ontwikkel wat die opkoms van meervoudige weerstandsmeganismes probeer oorkom. Hierdie meganismes word deur puntmutasies binne die ensiem veroorsaak. Die onlangse opkoms van weerstand in 2015, wat deur die EGFR-C797S-mutasie veroorsaak is, het 'n uitdagende terapeutiese poging aangewend. Geen groot deurbrake om hierdie klinies relevante mutantvariant te oorkom is nie tot op datum bereik nie. Die navorsingspogings in hierdie proefskrif fokus op die ontwikkeling van nuwe terapeutiese strategieë om hierdie weerstandsmutasie te oorkom.

Die sintese van hoe-kragtige omkeerbare en onomkeerbare EGFR-inhibeerders was die eerste van hierdie benaderings. Hierdie werk was gebaseer op 'n voorheen geïdentifiseerde suksesvolle inhibeerder. Verskeie glikkol-funksionaliteite, wat by strategiese posisies ingevoer is om gunstige omkeerbare interaksies binne die aktiewe plek van die ensiem voor te stel, was geïnspireer van dubbele EGFR/HER2-inhibeerders. Verdere studies was onderneem oor die gevolge van derivatisering en verwisseling van ander strukturele elemente, soos die oplosbare groep en akrielamiedeletrofil.

Ons tweede strategie, wat afwyk van tradisionele inhibisietodes, het gesoek na die herstelling van onomkeerbare inhibisie deur die EGFR katalitiese lysine 745 residu met nuwe elektrofile. Om kovalente bindingsvorming te veroorsaak was inhibeerders met 1,4-dikarbonyl-koppe gesintetiseer wat Paal-Knorr-pyrrolvorming kon ondergaan met die Lys745 residu. Verder is potensiële omkeerbare en onomkeerbare interaksies met die katalitiese lysienresidu ondersoek met inhibeerders wat nitril- en reaktiewe sulfonielfluorieddele bevat.

Laastens het ons kovalente binding vorming nagestreef met die gemuteerde serine 797 residu - die hoofmeganisme van weerstand in die EGFR-C797S mutant variant. Inhibeerders wat 'n verskeidenheid reaktiewe seriene-teiken-elektrofile bevat, insluitende boron-, fosfor-, swael- en fluoorbevattende hoofkoppies, was gesintetiseer om dit te bereik.

Bemoedigende resultate was opgelewer in die biochemiese en sellulêre evaluering van die 48 finale inhibeerders, wat tydens hierdie navorsingsprojek gesintetiseer is. Dit sluit in pikomolêre IC_{50} waardes vir biochemiese inhibisie van die klinies relevante EGFR-L858R/T790M/C797S triple mutant variant van die ensiem deur die gesintetiseerde omkeerbare inhibeerders. Hierdie resultate is egter ongelukkig nie weerspieël in die sel-gebaseerde toetse nie.

Daarbenewens het verskeie inhibeerders wat die konvensionele akrylamiedelektrofil bevat, lae nanomolêre en pikomolêre reeks aktiwiteit vertoon teen die EGFR-L858R/T790M dubbelmutant in 'n onderskeidelik biochemiese en sellulêre omgewing. Hierdie resultate het die onomkeerbare inhibisie van die ensiem aangedui. 'N Kristalstruktuur was verkry van die mees effektiewe onomkeerbare inhibeerder gekompleks saam met die cSrc-T338M/S345C-surrogaat. Hierdie struktuur het die kovalente bindingvorming met die Cys797-residu bevestig en ons postulate oor die waargenome tendense in aktiwiteit ondersteun. Laastens was verskeie van die mees suksesvolle inhibeerders geëvalueer en getoon dat dit 'n effektiewe dubbele EGFR/HER2 inhibeerder is.

'n Omkeerbare bindingsmodus deur die gebruik van kovalente massaspektrometrie-eksperimente was bewys vir alle inhibeerders wat die katalitiese Lys745 en mutant Ser797 residu geteiken het. Verskeie funksionele groepe is egter geïdentifiseer vir die vermoë om gunstige interaksies met hierdie residue te skep, soos wat weerspieël was in hul effektiewe biochemiese en sellulêre aktiwiteitsprofile teen die EGFR-C797S mutant. Dit sluit in die 1,4-dialkeen-, boronester-, vinielsulfonylfluoried-, imidasool- en nitril-draende verbindings.

Op pad na die sintese van hierdie inhibeerders en hul onderskeie elektrofile is nuwe sintetiese metodologie gevestig. Hoogtepunte sluit in die ontwikkeling van 'n nuwe en gerieflike eenpot-installasie van 'n asetamied- of propanamiedimidasool funksionele groep na anilienderivate, wat gebruik maak van 1,1'-karbonyldiimidasool en die ooreenstemmende α -, β -halo-karboksielsuur. Geoptimaliseerde metodes was afgelei vir toekomstige toegang tot die verskeidenheid elektrofile en heterosikliese strukture wat deur hierdie navorsingsprojek gesintetiseer word.

Acknowledgements

“At this time, organic chemistry can drive one completely crazy. It seems to me like a primeval tropical jungle, full of the most remarkable things, an amazing thicket, without escape or end, into which one would not dare to enter.”

- Friedrich Wöhler

The above quote by the pioneering chemist Frederick Wöhler captures and describes the scientific field which has become such a huge part of my life – synthetic organic and medicinal chemistry. To this day I am still in awe of how such small, highly-structured clusters of atoms can act as modern medicines – saving and prolonging human lives, alleviating pain and suffering and even being able to control cells, tissues and organs millions of times their size. Every day I feel truly blessed and grateful to be pursuing a passion that I love and appreciate so deeply – immersing myself in the understanding of this phenomena.

Earning a PhD is a gruelling task which is won through determination, grit and hard work. Throughout this journey I have been fortunate enough to be supported by so many people who have believed in me and played a significant part - big or small. First, I must thank my supervisor Prof. Willem van Otterlo. I can still remember the day I came into your office and said I was interested in taking on a medchem project in your research group. Many years later I can say it was one of the best decisions I have made. I have grown exponentially under your supervision as both an organic chemist and person. Thank you for being a mentor, a friend and a fair supervisor throughout my postgraduate career. Also, to my co-supervisor Dr. Stephen Pelly, while you had to leave to take up a position in Atlanta, I am still grateful for all the medicinal chemistry and molecular modelling knowledge you imparted to me while you were here. I look forward to new adventures and beer brewing escapades with you during my coming post-doc at Emory University.

Throughout my life I have been surrounded by incredibly strong women that have raised and guided me to becoming the man I am today. The person that has played the biggest role in this is my mother, Elschen. Thank you for every single sacrifice you have had to make in your life to make sure I was alright. Thank you for being the person that has always believed in me, always given and never asked for anything in return and always loved me no matter what. Through all the twists and turns of my life you have always been a rock of support. I love you and nothing I could ever do could repay the love and generosity you have always shown me.

To my beautiful girlfriend Alet, I still remember the first time I laid my eyes on you while standing in the lab. If I wasn't given the opportunity to study, I would never have met you - something I am grateful for every day. Thank you for teaching me to look after myself, for your kind nature and for being my shelter from the storm. Your love has transformed me into the person I was meant to be. I am so excited for our future together.

My dear sister Loreen, I cherish our relationship and all we have been through and overcome together. Both you and Ryan are a pillar of strength in our family. Thank you for the sacrifices you had to make to keep us afloat financially and for all you have given me over the years. I would not be at this point if it weren't for your generosity. I look forward to the day when all our children will be playing on the beach together.

To my father, Derek, thank you for all the sage advice you have given me throughout my life. From being a great dad and raising two children with trips to the Kruger park, Namibia and Kariega you have become one of my best friends. I always look forward to our suppers filled with treasured conversations and lots of whiskey. Here's to many more.

I must extend my heartfelt gratitude to the other academic staff of the group of organic and medicinal chemistry. I would like to thank Prof. Gareth Arnott in particular. From lecturing an unruly undergraduate student all the way to discussing complex reaction mechanisms, you have always had an open-door policy and shared your knowledge freely. To Prof. Ivan Green, thank you for letting me inhabit your office and share your space for the past year. I will look back fondly on the conversations we had and for all the chemistry and life advice you offered me. You have been an incredibly valuable asset to the department, and I wish you well in your retirement.

I would like to especially thank Prof. Daniel Rauh for hosting me on my research visit to Dortmund in 2016. I absolutely thrived and learnt an incredible amount working in the multi-disciplinary environment you have created for your students. I am fortunate enough to say I made lifelong friends during my time abroad. Amongst others, I would like to thank Jonas Lategahn, Mike Buhrman, Tobias Grabe, Steven Smith, Marina Keul, Denise Dos Santos and Hannah Tumbrinck for making my stay such a memorable and pleasurable one.

To my best friends Jacques Vogeli, Leon Jacobs and Derik Wilbers - thank you for all the support throughout the years. Even if that support came in the form of far too many beers - and led to a very unsupportive hangover - I will always treasure the crazy memories. I place immense value in our friendships and even though our careers and lives might send us across the globe and separate us, I know we will always remain friends.

A very big thank you to the technical staff of the department, who over the years have become my friends. Your daily work and efforts keep the cogs of the machine turning. Raymond, Debbie, Maxwell and Shafiek, thank you for your constant diligence and positive personalities, especially during the undergrad practical experiments. Also, thank you to Mr de Jongh in his retirement who served the De Beers building for so long and to Mary in her passing - you are sorely missed every day in the building, may you rest in peace.

I would like to thank the Central Analytical Facility at Stellenbosch University for the services they have provided throughout the years. Elsa Malherbe and Dr. Jaco Brand, thank you for your patience, constant willingness to help and positive words in the NMR department. A special thank you to Dr. Marietjie Stander who helped me secure funding in 2017 and for running all mass spectrometry related experiments.

Without financial support I would not have been able to undertake my postgraduate studies and be where I am today. I would like to thank my tannie Ingrid for never giving up on her belief in me, even in the most dire and dark of times. Thank you for your financial support when I needed it most and for being a rock of support for our family. I would like to thank Dr. Rehana Malgus-Enus for her generosity in helping me secure funding in 2018. Lastly, I would like to thank the Medical Research Council of South Africa for funding my research visit overseas, the National Research Foundation for the various bursaries throughout my postgraduate career and Stellenbosch University for their facilities, funding during my honour's year and the merit bursary.

Table of Contents

Chapter 1 - The Burden of Cancer, Cancer Biology and Targeted Treatment

1.1 The Burden of Cancer	1
1.1.1 Incidence and Mortality	1
1.1.2 Cancer Disparity	2
1.1.3 Economic Impact	3
1.1.4 Prevention	3
1.2 Cancer Biology: Enabling Characteristics and Hallmarks	4
1.2.1 Cancer Genomics and Epigenetics.....	5
1.2.1.1 The Complexity of Cancer Genomics.....	5
1.2.1.2 Driver and Passenger Mutations.....	5
1.2.1.3 Oncogenes.....	7
1.2.1.4 Tumour Suppressor Genes.....	7
1.2.1.5 DNA Repair Genes.....	8
1.2.1.6 Cancer Epigenetics	8
1.2.2 Tumour Microenvironment	8
1.2.2.1 Inflammation.....	8
1.2.2.2 Angiogenesis	9
1.2.2.3 Metastasis.....	9
1.2.3 Emerging Hallmarks	10
1.2.3.1 Cellular Metabolism	11
1.2.3.2 Immune Evasion.....	12
1.3 Chemo- and Targeted Therapy of Cancer	12
1.3.1 History.....	12
1.3.1.1 Chemotherapy	12
1.3.1.2 The Targeted Therapy Revolution	13
1.3.2 Modern Targeted Therapy.....	14
1.3.2.1 Industry Overview	14
1.3.2.2 Challenges and Limitations	15
1.3.2.3 Targeting the Hallmarks	16
1.4 References	17

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

2.1 Epidermal Growth Factor Receptor	25
2.1.1 Discovery	25
2.1.2 Receptor Structure and Ligand Binding.....	27
2.1.3 The ErbB Signalling Network	29
2.1.3.1 Input Layer.....	29
2.1.3.2 Signal-processing Layer.....	30
2.1.3.3 Output Layer.....	30
2.1.4 Oncogenic Activation of EGFR and Cancer	31
2.1.4.1 EGFR Overexpression.....	31
2.1.4.2 Overproduction of Ligands	32
2.1.4.3 EGFR Cross-talk and Heterodimerization	32
2.1.4.4 Defective EGFR Downregulation	32
2.1.4.5 Activating EGFR Mutations.....	33
2.2 Protein Kinases	34
2.2.1 Origins of Reversible Protein Phosphorylation.....	34
2.2.2 The Human Kinome	35
2.2.3 The Mechanism of Protein Phosphorylation and the Kinase Active Site	35
2.3 Protein Kinase Inhibitors	37
2.3.1 Overview.....	37
2.3.2 Types of Kinase Inhibitors.....	38
2.3.2.1 Type I and II Inhibitors	38
2.3.2.2 Type III and IV Inhibitors	38
2.3.2.3 Type V Inhibitors.....	39
2.3.2.4 Type VI Inhibitors.....	39
2.4 EGFR as Therapeutic Target.....	40
2.4.1 1 st Generation Inhibitors.....	41
2.4.2 2 nd Generation Inhibitors.....	42
2.4.3 3 rd Generation Inhibitors	44
2.4.4 4 th Generation Inhibitors and the Future of EGFR Inhibition	47
2.5 References	48

Chapter 3 – Overview of Aims, Objectives and General Practices

3.1 Introduction	52
3.2 Chapter 4: Lead Optimization of Pyrrolopyrimidine-derived EGFR Inhibitors.....	52
3.3 Chapter 5: 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745	53
3.4 Chapter 6: Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745	55
3.5 Chapter 7: Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797	56
3.6 Concluding Remarks.....	57
3.7 General Practices.....	58
3.7.1 Synthesis.....	58
3.7.1.1 Solvents and Reagents	58
3.7.1.2 Chromatography and Purification	58
3.7.1.3 Spectroscopic Characterization and Physical Properties.....	59
3.7.1.4 Glassware, Inert Conditions and Temperature Control	59
3.7.2 Biochemical and Cellular Evaluation.....	60
3.7.2.1 Protein Expression and Purification	60
3.7.2.2 Construct Design of EGFR-L858R/T790M/C797S	60
3.7.2.3 Activity-based Assay for IC ₅₀ Determination	60
3.7.2.4 Cell Culture.....	61
3.7.2.5 Viability Assay (EC ₅₀ Determination).....	61
3.7.2.6 Covalent Mass Spectrometry Experiments	61
3.8 References	62

Chapter 4 – Lead Optimization of Pyrrolopyrimidine-derived EGFR Inhibitors

4.1 Introduction	63
4.2 Lead Compound Identification and Optimization	63
4.3 Inspiration for Dual EGFR/HER2 Inhibitors	66
4.4 Objectives and Target Synthetic Combinatorial Library.....	67
4.5 Combinatorial Library Synthesis	69
4.5.1 Pyrrolopyrimidine Scaffold Synthesis	69
4.5.2 Synthesis of Series A and B.....	70
4.5.3 Synthesis of Series C and D	75

4.5.4 Synthesis of Series E and F.....	78
4.5.5 Synthesis of Series G and H.....	80
4.6 Biochemical and Cellular Evaluation	82
4.6.1 Comparison of Reversible Inhibitors	83
4.6.2 Comparison of Potentially Irreversible Inhibitors	86
4.6.3 Further Evaluation of Optimal Irreversible Inhibitors	89
4.7 Crystal Structure of Compound 77	90
4.8 Conclusions.....	92
4.9 Ongoing and Future Work.....	94
4.10 Supplementary Information	96
4.11 References	135
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745	
5.1 Introduction	137
5.2 Targeting the EGFR Catalytic Lysine Residue	138
5.2.1 Catalytic Control in EGFR and Kinase Regulatory Mechanisms	138
5.2.2 Lysine-trapping Covalent Inhibitors.....	140
5.3 Objectives and Target Synthetic Library	143
5.4 Library Synthesis.....	144
5.4.1 Heterocyclic Driving Group Synthesis.....	144
5.4.2 1,4-Diketone Compound Synthesis	146
5.4.3 1,4-Dicarboxylic Acid and 1,4-Diester Compound Synthesis.....	148
5.4.4 Attempted 1,4-Dialdehyde Compound Synthesis	151
5.4.4.1 Reduction of 1,4-Diester Containing Compounds	151
5.4.4.2 Oxidation of 1,4-Diol-containing Compounds	152
5.4.4.3 Ozonolysis and Wacker Oxidation of 1,4-Diene Containing Compounds	155
5.5 Biochemical and Cellular Evaluation	158
5.6 Conclusions.....	161
5.7 Future Work.....	162
5.8 Supplementary Information	165
5.9 References	177

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745	
6.1 Introduction	180
6.2 Electrophile and Heterocyclic Scaffold Selection	181
6.2.1 Additional Lysine-trapping Electrophiles	181
6.2.1.1 Sulfonyl Fluorides	181
6.2.1.2 Michael Acceptors	182
6.2.1.3 Nitriles and α -Haloketones	183
6.2.2 Heterocyclic Scaffolds Exploiting Catalytic Lysine Interactions	187
6.3 Objectives and Target Synthetic Library	187
6.4 Library Synthesis.....	188
6.4.1 Osimertinib-derived Scaffold Synthesis	188
6.4.2 Electrophile Synthesis and Derivatisation	197
6.4.2.1 Michael Acceptors.....	197
6.4.2.2 Nitriles	199
6.4.2.3 Sulfonyl Fluorides	200
6.4.2.4 α -, β -Haloketones.....	204
6.5 Biochemical and Cellular Evaluation	205
6.5.1 Evaluation of Intermediate Compounds	205
6.5.2 Evaluation of Intermediate Compounds	208
6.6 Conclusions.....	209
6.7 Future Work.....	211
6.8 Supplementary Information	215
6.9 References	233
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797	
7.1 Introduction	235
7.2 Targeting Serine Residues	236
7.2.1 The Catalytic Triad.....	236
7.2.2 Serine-targeting Covalent Inhibitors	238
7.2.2.1 Previously Discussed Electrophiles	238
7.2.2.2 Fluorophosphonates	238

7.2.2.3 Boronic Acids	239
7.3 Objectives and Target Synthetic Library	240
7.4 Library Synthesis.....	242
7.4.1 Osimertinib-derived Scaffold Synthesis	242
7.4.2 Model Compounds	243
7.4.3 Electrophile Synthesis and Derivatisation	244
7.4.3.1 Ethenesulfonamide	244
7.4.3.2 Sulfonyl Fluorides	246
7.4.3.3 Phosphonates and Fluorophosphonates	248
7.4.3.4 Boronic Esters and Acids.....	252
7.4.3.5 Trifluoromethyl ketones.....	258
7.4.3.6 α -, β -Haloketones.....	258
7.4.3.7 Imidazole Compounds	260
7.4.3.8 Nitriles	264
7.4.3.9 Dual Warheads	264
7.5 Biochemical and Cellular Evaluation	272
7.6 Conclusions.....	281
7.7 Future Work.....	284
7.8 Supplementary Information	290
7.9 References	314
Chapter 8 – Project Summary and Concluding Remarks	
8.1 The Future of EGFR as Therapeutic Target	317
8.2 Summary, Outcomes and Future Applications of this Project.....	318
8.3 References	320

Chapter 1

The Burden of Cancer, Cancer Biology and Targeted Treatment

Abstract

In this chapter, the burden of cancer and its effects on all of mankind will be discussed. Emphasis will be placed on its human, societal and economic costs and preventative measures available. An understanding of cancer biology, and the inherent events and characteristics that confer selective growth and survival advantages to its cells, is essential to deciphering the pathogenesis of the disease. These enabling and sustaining capabilities, known as the hallmarks of cancer, will be reviewed. Finally, an overview of chemo- and targeted-therapy, limitations and challenges in the field, and how the hallmarks have influenced cancer drug development, will be reflected on.

1.1 The Burden of Cancer

Cancer is the collective name given to an extensive group of diseases which is mainly characterised by the uncontrolled growth of abnormal cells that either ignore or evade biological processes or display enhanced biological processes. This unregulated cellular growth results in the rapid division of these aberrant cells, leading to the formation of malignant neoplasm or tumours.¹ These cells have the ability to detach from primary tumour sites and penetrate surrounding tissue, entering blood or lymphatic vessels and allowing for circulation via the vascular or lymphatic system. This results in the invasion of adjoining body parts or vital organs and the formation of secondary tumour sites – a process known as metastasis.² Cancer can evolve from nearly all cells and can develop in and spread to all parts of the human body. It is an incredibly complex disease group which has many anatomic and molecular subtypes, each of which require tailored management strategies.

1.1.1 Incidence and Mortality

Cancer is a costly burden and first and foremost is the cost of human life. Cancer is the second leading cause of death worldwide, claiming an estimated 9 million lives and accounting for 1 in 6 deaths in 2016.³ Of the estimated 57 million deaths globally in 2016, deaths occurring due to noncommunicable diseases (NCD) accounted for a staggering 41 million (71%) of these.³ The clear majority of these deaths can be attributed to four main culprits, namely: cardiovascular disease (17.9 million, 44% of NCD deaths), cancer (9 million, 22%), chronic respiratory disease (3.8 million, 9%) and diabetes (1.6 million, 4%).³ The five leading behavioural risks for cancer, and the shared risk factors for the four main NCDs, are: not maintaining a healthy body weight, lack of regular exercise, inadequate intake of fruits and vegetables, use of tobacco products and increased alcohol consumption.^{4, 5} Tobacco use is the single greatest avoidable risk factor for cancer and is responsible for 22% of all cancer related deaths.⁵ While globally, the risk of death from one of the four major NCDs has fallen from 22% in 2000 to 18.3% in 2016, adults in lower- and middle-income countries still face nearly double the risk (22%) of that of adults in high-income countries (12%).⁶ Locally, the continental average risk for Africa was estimated at 20.6%, with South Africa (26.2%) placed 5th highest amongst the 47 countries taken into consideration.⁶ Thus, these noncommunicable diseases constitute a major obstacle to human development and well-being.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

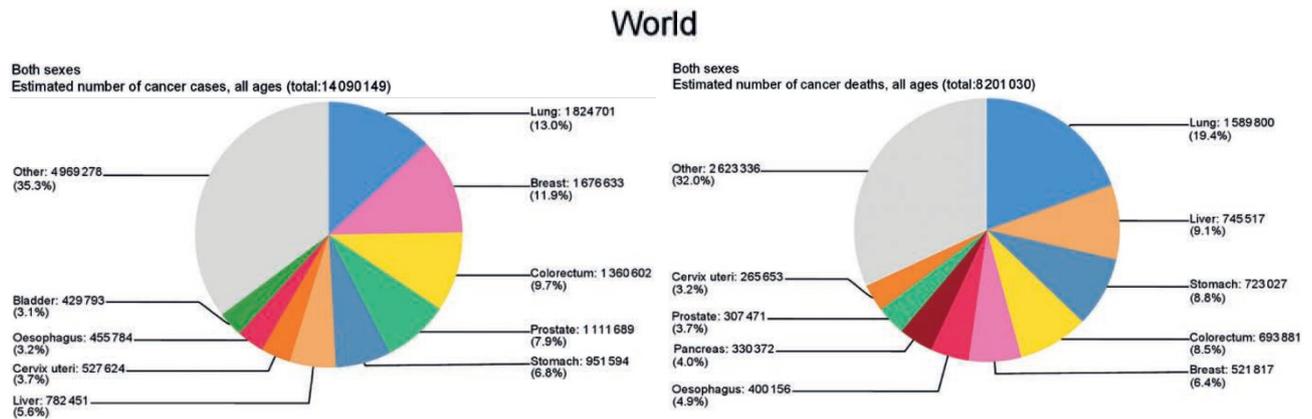


Figure 1.1: The estimated cancer incidence (left) and mortality (right) by major site, in both sexes combined in 2012. Figure reproduced from Stewart et al.⁷

According to the World Cancer Report 2014, the incidence of cancer has increased from 12.7 million in 2008 to 14.1 million in 2012.⁷ As illustrated in **Figure 1.1**, the five most common sites for cancer diagnosed in both men and woman combined were the lung (13.0% of total cases), breast (11.9%), colorectal (9.7%), prostate (7.9%) and stomach (6.8%) – constituting half of the overall global cancer burden.^{7, 8} A similar pattern is observed for the five deadliest cancers, with cancer of the lung (19.4% of total deaths), liver (9.1%), stomach (8.8%), colorectum (8.5%) and breast (6.4%) having the highest mortality rates (**Figure 1.1**).⁸ Over the next two decades the incidence of cancer is set to rise by a further 75%, bringing the expected number of cancer cases close to 25 million in 2035.⁷ The projected escalation in cancer will undoubtedly have the greatest impact in low- and middle-income countries.

1.1.2 Cancer Disparity

Whilst the harrowing experience of being diagnosed with cancer is a truly universal and unbiased one, patients find their paths post-diagnosis to be uniquely diverse. The personal experiences of these patients frequently reflect the disparity and inequality found globally, and the future of a patient depends largely on where they find themselves in the world. More than 60% of the world's cancer incidents, and approximately 70% of deaths from cancer, occur in the regions of Africa, Asia and Central and South America.⁷ This can be attributed to health care systems that are ill equipped to cope with the burden of cancer. Worldwide, only 1 in 5 low- to middle-income nations have the requisite data to drive cancer policy.⁹ Consequently, cancer is typically diagnosed at more advanced stages of the disease and access to diagnosis, effective treatment or palliative care are often inaccessible. In 2017, it was found that only 26% of low-income nations reported having publicly available pathology and treatment services, in stark contrast to over 90% for that of high-income countries.¹⁰

Many low- to middle-income countries find themselves having to contend with cancers from two vastly different worlds. Cancers associated with poverty, such as those attributable to the infectious agents of *helicobacter pylori*, Epstein-Barr virus, hepatitis and human papilloma virus (HPV) remain a problem. In 2012, infection-attributable cancers made up 15% of global cancer cases, of which more than 66% percent of these occurred in low-income countries and less than 5% from Australia, Canada, New Zealand, the United States and select countries in Western and Northern Europe combined.¹¹ Infection-related cancers accounted for 25% of total cancer cases in developing countries, and more than 50% in sub-Saharan Africa.¹¹

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

Additionally, cancers associated with the world of plenty are becoming increasingly prevalent. This can be ascribed to the adoption of Western lifestyles – lack of physical activity, consumption of unhealthy, processed foods and alcohol and increased use of tobacco. It has been estimated that nearly 80% of the 1 billion smokers worldwide now reside in low- to middle-income nations.⁵ The rising incidence of cancer and other NCDs will place immense strain on the already struggling healthcare systems of developing countries, leading to the divide between experiences of individual cancer patients to broaden.

1.1.3 Economic Impact

Treatment for an increasing number of cancer patients also has an escalating economic impact. In 2010, the total annual cost of cancer, with regards to prevention, treatment and the economic value of lives lost and disability caused, was estimated to be \$1.16 trillion.⁷ Further analysis, which adds the long-term costs to patients and their family members, estimate the annual global cost at \$2.5 trillion.⁷ High-income countries spend between 5-10 times more on cancer management and prevention on a per capita basis than low- and middle-income nations.¹² Whilst many assessments have been undertaken on the cost of cancer in these high-income regions, a recent study estimated the value of productivity lost in 2012 due to cancer mortality in Brazil, Russia, India, China and South Africa – also known as the BRICS countries.¹³ Utilising data from GLOBOCAN 2012, the total cost of lost productivity from cancer deaths in all five BRICS countries was valued at \$46.3 billion, constituting 0.33% of their combined gross domestic product and equating to \$45926 per premature cancer mortality.¹³ It was found that China suffered the greatest loss in years of productive life (5.9 million years) and total productivity (\$28 billion), while South Africa had the highest cost per cancer death (\$101105) – more than five times the cost estimated in India (\$19 691).¹³ Factors such as ageing, population growth and the spiralling costs of cancer treatment will place increasing demands on the health-care budgets of wealthy and impoverished nations alike.

1.1.4 Prevention

Given the overwhelmingly heavy burden of cancer globally, and the projected impact on developing countries particularly, prevention is fundamental to diminishing and reversing this burden. It is estimated that 30 – 50% of all cancer cases are preventable by avoiding or modifying the five leading behavioural and dietary risks.⁴ In low- to middle-income regions of the world, it has been estimated that an investment of \$11.4 billion in a set of core prevention strategies could lead to savings of up to \$100 billion in future cancer treatment costs.¹⁴ Approaches with the greatest potential for clinical and economic results include tobacco and obesity control, vaccines, early detection and the creation of cancer registries.¹⁴ Through the implementation of vaccination against HPV and hepatitis B viruses alone, it is estimated that 1 million cancer cases can be prevented annually.¹¹ By adopting similar strategies for cancer control in the BRICS (Brazil, Russia, India, China and South Africa) nations, 41% of the world's population will increase their chances of a cancer-free life.¹³ Elucidating the causes and creating effective prevention strategies will thus be essential components of combating the growing cancer burden in the future.

Parallel to efforts engaged in the causes and prevention of cancer, exceptional progress has been made in the understanding of the molecular and biological events that transform a normal functioning cell into a malignant growth, capable of replicating and spreading throughout our bodies.¹⁵

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

These advances have in turn led to new avenues in early detection and prevention, clues relating to causation and the classification of different types of cancer.¹⁶ This integrative and complementary approach is imperative to translating knowledge about the mechanisms of carcinogenesis into freeing as many people as possible from the threat of this disease.

1.2 Cancer Biology: Enabling Characteristics and Hallmarks

A grasp and comprehension of cancer biology is pivotal to the development of rationally designed therapeutics and preventive recourse. At the dawn of the new millennium, the Human Genome Project revealed the specification of the human genome, providing a comprehensive view of a “normal” human sequence and its components.¹⁷ In the following years, the investigation of how cancer cells differ from normal cells has yielded an invaluable amount of relevant knowledge to the field – with acquisition of this knowledge proceeding exponentially faster in a post-genomic era. The hallmarks of cancer, shown in **Figure 1.2**, were first laid out in an influential review in 2000 by Hanahan and Weinberg, providing an enumeration of the traits and defining characteristics of this complex disease.¹⁸ These principles were further elaborated in an update published in 2011, proposing new hallmarks and mechanisms of cancer.¹⁵ These hallmarks and the pathogenesis of cancer will be discussed within this section, placing emphasis on the enabling characteristics of somatic mutations and altered gene expression within the cancer genome. These mutations may occur through epigenetic change and mediate selective growth advantages conferred by the hallmark capabilities. In parallel, the role of the tumour microenvironment in facilitating tumorigenesis through inflammation, growth of new blood vessels and the spread and growth of cancer cells in metastatic sites will also be discussed.

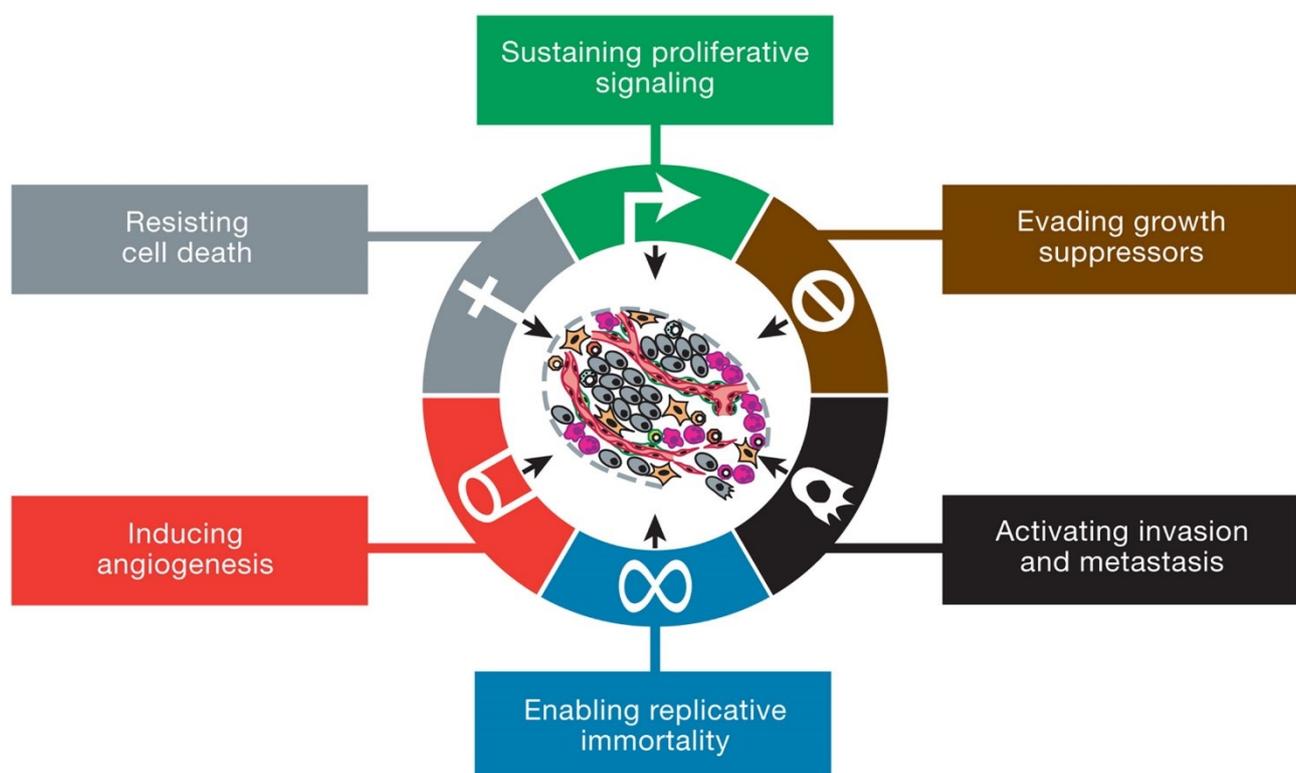


Figure 1.2: The six hallmark capabilities as proposed by and reproduced from Hanahan and Weinberg.¹⁵

1.2.1 Cancer Genomics and Epigenetics

1.2.1.1 *The Complexity of Cancer Genomics*

Cancer is a disease of the genome and all cancers harbour genetic mutations within their own genomes. The biological complexity of the disease stems from several factors regarding these mutations, the first being the numerous types of cancer-causing mutations.¹⁹ The simplest of these mutations can be as a result of the insertion, deletion or substitution of only one nucleotide that affects single DNA bases. Consequently, segments of chromosomes can undergo gene amplification, deletion or inversion of gene sequences.^{20, 21} Translocation of chromosomes may also occur, most notably exemplified in the Philadelphia translocation which results from the joining of chromosomes 9 and 22, a mutation prevalent in over 95% of patients suffering from chronic myeloid leukaemia (CML).²² Viral carcinogenesis can also occur through the acquisition of DNA from viruses such as HPV, hepatitis and the Epstein-Barr virus.²³

In addition to the extensive mutational aetiology of cancer, the sheer number of different mutations present in tumours adds to their complexity. Using whole cancer genome sequencing, the mutational spectrum of cancer cells has been elucidated, revealing a substantial number of putative somatic mutations in every cancer genome. A total of 22910 somatic base substitutions were observed in the first small cell lung cancer genome sequenced, and 33345 in the first sequencing of the melanoma genome.^{24, 25} Multiple mutation signatures could be traced to repeated exposure to carcinogens such as ultraviolet radiation and chemicals found in tobacco smoke, and their proclivity towards defective DNA repair pathways.

Lastly, tumour heterogeneity, or the composition of different cancer cell subpopulations in tumour masses bearing related but distinct mutation profiles, can be highly problematic. Over time and through the process of metastasis, the mutational spectrum of lesions in a single patient can evolve. This evolution can have significant ramifications, the most serious being the acquisition of drug resistance.²⁶ Tumour heterogeneity promotes the emergence of more-competitive cancer sub-clones that are insensitive to therapy, as they outlive those clones that are not. Genomic instability can also alter the pharmacogenomic pathway or the drug receptor site of the therapeutic target, resulting in an adverse drug reaction or rendering the treatment ineffective.²⁷ This is particularly relevant in the treatment of non-small cell lung cancer (NSCLC) with epidermal growth factor receptor (EGFR) as the therapeutic target. In the last 15 years alone, the two single point mutations T790M and C797S have mediated resistance to a number of FDA approved blockbuster drugs including gefitinib, erlotinib and osimertinib.^{28, 29} The aforementioned case remains the exception as the development of cancer can be seen as a multistep, multigenic process arising not from a single polymorphism in a single gene, but rather the accumulation of a number mutations over time.

1.2.1.2 *Driver and Passenger Mutations*

Mutations have profound effects on the biological processes that drive cancer cell growth and differentiation. While all cancers arise from somatic changes in the DNA of cancer cells, not all the somatic abnormalities acquired in a cancer genome are implicated in the development of the cancer. Mutations that confer selective growth advantages to cancer cells are known as driver mutations.³⁰ Driver mutations are causally implicated in oncogenesis and are found within the subset of genes known as cancer genes.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

While they are not required for preservation of the final cancer, they have been positively selected at some point during the evolution of the cancer.³¹ The remaining majority of somatic mutations with no biological consequence are called passenger mutations. Passenger mutations are without functional significance as they do not confer clonal growth advantages and do not contribute to the incidence of cancer. These mutations often occur during mitosis and are carried along during clonal expansion so that they are found in all cells of the final cancer.³⁰

For more than 25 years, a central aim of cancer research has been the identification of cancer genes which harbour driver mutations.³¹ Recently, through the use of cancer genome sequencing, a comprehensive analysis of 9423 tumour exomes was undertaken. This combined all 33 of The Cancer Genome Atlas categorized cancer types and made use of 26 different computational tools to provide the most up to date catalogue of driver genes and mutations.^{16, 32} The study was able to identify 299 cancer driver genes and over 3400 putative driver mutations.³² The research also found that 57% of the tumours analysed harboured potentially actionable oncogenic events, a finding that can serve as a foundation for future clinical and therapeutic endeavours.³² Driver genes causing alterations which result in carcinogenesis can be divided into three broad groups, namely: proto-oncogenes and oncogenes, tumour suppressor genes and DNA repair genes.³¹ These three groups can be further classified into 12 signalling pathways that control three central cellular processes: cell fate, cell survival and genome maintenance as shown below in **Figure 1.3**.³¹

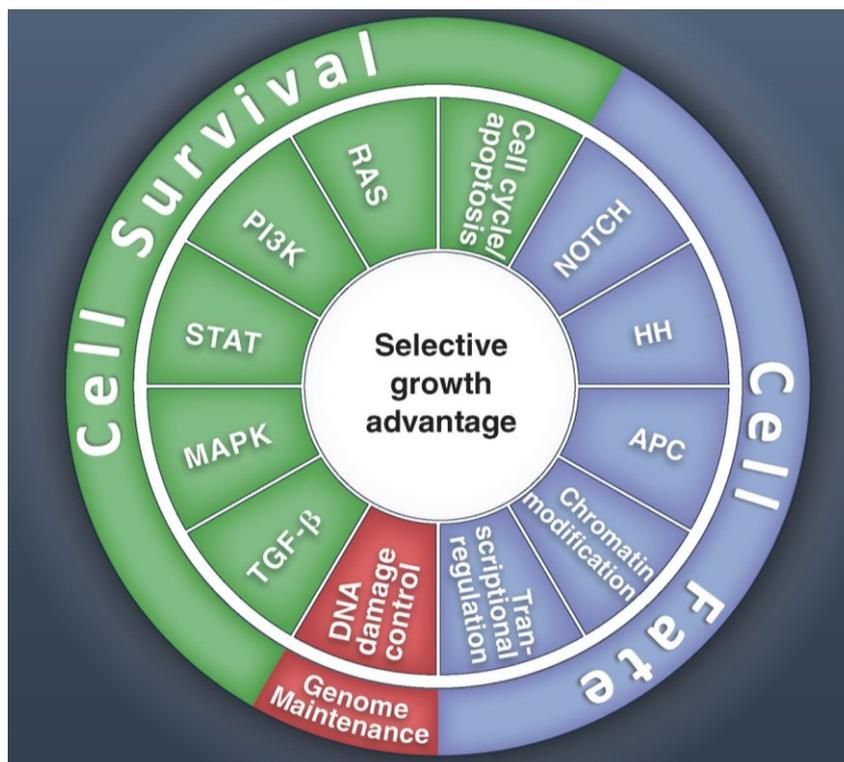


Figure 1.3: Selective growth advantages (inner ring) is associated with driver gene mutations in one or more of 12 cell signalling pathways (middle ring). These pathways are responsible for the regulation of three core cellular processes (outer ring). Figure reproduced from Vogelstein et al.³¹

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

1.2.1.3 Oncogenes

In normal cells, proto-oncogenes code for the proteins that stimulate cell division, inhibit cell differentiation and halt cell death through a series of steps called signal transduction cascades and pathways. Upon acquiring an activating mutation, a proto-oncogene becomes a tumour inducing-agent known as an oncogene, leading to upregulation and over expression of these proteins.³³ This activation is induced by three main mechanisms, namely: mutations within regulatory regions changing the structure of the encoded protein to enhance its transforming activity, gene amplification or duplication leading to an increase in protein concentration and expression and chromosomal rearrangements including inversions and translocations.^{34, 35} Mutations and translocations can occur as initiating events or during tumour progression, whilst amplification is mainly associated with tumour progression.³³ All three mechanisms confer a growth advantage and increased cell survival to cancer cells through an increase in or deregulation of protein expression.³⁶ Oncogenes are almost always dominant as only a single allele mutation is required for the gene to possess a new molecular function or pattern of gene expression. This is referred to as a “gain of function” mutation.³⁷

The products of oncogenes can be divided into six categories: transcription factors, chromatin remodelers, apoptosis regulators, signal transducers, growth factors and growth factor receptors – being implicated in the majority of the hallmark capabilities of cancer.³³ Notable examples of oncogenes include the RAS gene superfamily, where single point mutations typically produce constitutive constant activation of the RAS signalling pathway. About 25% of all human tumours contain an activating mutant isoform of RAS genes.³⁸ MYC oncoproteins belong to a family of super transcription factors that regulate the transcription of at least 15% of the entire human genome. The MYC family of oncogenes are deregulated in over 50% of human cancers, being often associated with adverse prognosis and poor patient survival rate.³⁹ Oncogenic proteins in cancer cells remain therapeutic targets for cancer treatment, as their targeted inactivation results in sustained loss of neoplastic growth.⁴⁰ Growth factors and growth factor receptors form an integral part of this research project and their signalling pathways, relevance as therapeutic targets and current therapies will be comprehensively reviewed in the following chapter.

1.2.1.4 Tumour Suppressor Genes

Tumour suppressor genes, or anti-oncogenes, are distinct from oncogenes. Where proto-oncogenes can be seen as the accelerator in the vehicle that is the cell, tumour suppressor genes act as the brakes. Tumour suppressor genes code for the proteins that have a dampening or restraining effect on cell growth and differentiation and stimulate programmed cellular death.⁴¹ Unlike oncogenes, abnormalities within tumour suppressor genes are typically recessive and follow the “two hit hypothesis”.⁴² This requires that both alleles which code for a particular protein must be mutated for the effect to manifest and the trait to be expressed. This is referred to as a “loss of function” mutation.⁴¹ Tumour suppressor genes can be classified into three groups, namely: caretaker, gatekeeper and landscaper genes.⁴³ TP53 is an example of one of the most frequently mutated tumour suppressor genes, being implicated in more than 50% of human cancers.⁴⁴ Furthermore, TP53 encodes for at least 15 protein isoforms, known as the p53 protein isoforms, and has been described as “the guardian of the genome.”⁴⁵ These proteins provide genomic stability by binding to DNA and regulating gene expression to prevent mutations, whilst also enforcing a system of checks and balances to monitor cell division and apoptosis.⁴¹

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

1.2.1.5 DNA Repair Genes

The third group of genes associated with cancer are the DNA repair genes. Our cellular DNA is under continual attack from environmental agents such as radiation and chemicals, resulting in molecular lesions that cause structural damage. This may induce mutations in the cells genome, affecting the daughter cells survival after mitosis and altering or eliminating the cells ability to transcribe genes.⁴⁶ These toxic and mutagenic consequences are minimized by the action of DNA repair genes and the distinct repair pathways. These genes code for enzymes that repair point mutations and chromosomal aberrations before replication and chromosomal divide. Additionally, they are involved in cell-cycle checkpoints and, as a last resort, induce apoptosis if repair is not possible.⁴⁷ When these genes are mutated, the DNA damage response is silenced, leading to irreparable damage in cells which avoid programmed cellular death and undergo unregulated cellular growth. This results in the accumulation of subsequent mutations and ultimately in cancer.⁴⁸ Notable examples of DNA repair genes include BRCA, which encodes for protein products that are important for the repair of double strand breaks through homologous recombination. Mutations of these genes are typically implicated in breast cancer.⁴⁹

1.2.1.6 Cancer Epigenetics

Cancer epigenetics is the study of heritable modifications to the DNA of cancer cells that are not caused by genetic mutations and their biochemical pathways.⁵⁰ Epigenetic abnormalities participate with genetic mutations to cause dysregulation in gene silencing, playing a pivotal role in key biological processes which are fundamental to the genesis of cancer.⁵¹ These epigenetic alterations are regulated and brought about by DNA methylation, nucleosome remodelling by histone modifications and non-coding RNA-mediated targeting of various genes.⁵¹ The correlation between microRNAs (miRNAs), an abundant class of small non-protein-coding RNA molecules, and human cancers is especially well documented. When miRNAs are mutated or mis-expressed they are known as oncomirs.⁵² Oncomirs may act as both oncogenes - as overexpression of the gene results in uncontrolled cellular growth - and as tumour suppressors in normal cells, where under-expression of the gene is linked to cancerous growth.⁵³

With the recent advances in genomic technologies and the availability of an abundance of genomic information pertaining to cancer, the field of cancer genomics has shown a convergence in discovery science and clinical medicine. The successful translation of cancer genomics into diagnostics and therapeutics has reinforced the prospect of personalised cancer treatment.⁵⁴ Genomes of cancerous tumours will continue to be scrutinised and mined to further our understanding of cancer cells, to provide biomarkers as indicators of prognosis and treatment response and to identify future targets for the evaluation and implementation of new cancer drugs.

1.2.2 Tumour Microenvironment

1.2.2.1 Inflammation

Inflammation is an essential protective response that sporadically goes awry and becomes a major contributing factor in the pathogenesis of several chronic human diseases, including cancer. Since the discovery of leukocytes within neoplastic tissue in 1863 by Rudolf Virchow, the link between inflammation and cancer has become well established.⁵⁵ Whilst the genetic changes that occur within cancer cells allow for tumour initiation and promotion, this is not sufficient for the facilitation of tumour progression.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

Ancillary processes orchestrated by inflammatory cells within the tumour microenvironment are an indispensable participant and frequently play requisite roles in all stages of tumour development, from initiation and promotion to proliferation, invasion and metastasis.⁵⁶ Inflammatory responses enable and drive tumorigenesis through the supply of bioactive molecules to the tumour microenvironment such as growth factors for sustained proliferative signalling, survival factors limiting apoptosis and enzymes that expedite angiogenesis, invasion and metastasis.⁵⁷ As such, it is considered one of the enabling characteristics of cancer development.

1.2.2.2 Angiogenesis

Due to mutations that stimulate uncontrolled proliferation of cancer cells, tumours require substantially more nutrients and oxygen to sustain their growth and an enhanced ability to evacuate metabolic waste and carbon dioxide in comparison to normal tissues. This need is addressed by tumour-associated neovasculature or angiogenesis. During the development of an embryo, endothelial cells assemble into tubes and form blood vessels (vasculogenesis), which is followed by the sprouting of new vessels from these existing ones (angiogenesis) - a process which becomes largely quiescent in adulthood. However, during tumour progression, this “angiogenic switch” is unceasingly activated and remains on, resulting in the continuous sprouting of dilated and irregularly shaped tumour blood vessels that sustain expanding neoplastic tissue.⁵⁸ This “angiogenic switch” is controlled by counteracting factors that may either induce or oppose angiogenesis. Tumours secrete growth factors such as vascular endothelial growth factors (VEGF) that bind to receptors and induce capillary growth, and typically halt production of anti-VEGF enzymes such as protein kinase G (PKG) which prohibits angiogenesis in normal cells.¹⁵ Angiogenesis is induced in the early stages of tumour initiation and is fundamental to their characteristic large size. Angiogenesis also plays a role in tumour metastasis and has become one of the most relevant hallmarks of cancer in a therapeutic context.⁵⁹

1.2.2.3 Metastasis

Metastases result from the spread of cancer cells from a primary to a distant secondary tumour site and are responsible for more than 90% of all cancer-related deaths.² Metastasis, depicted in **Figure 1.4a**, is an intricate and multistep process where cancer cells migrate and flow through vastly different microenvironments such as the stroma, the vascular or lymphatic system and tissue at the secondary site. The success of this relies heavily on the physical, mechanical and molecular interactions of cancer cells with the micro- and external environment.⁶⁰ As mentioned previously, angiogenesis plays a pivotal role in the metastatic process, not only by providing nourishment through vascularisation, but also by providing an escape route by which cancer cells can enter the blood circulatory system.⁶¹ Investigation into the detachment of carcinoma cells have revealed a process resembling the epithelial-to-mesenchymal transition (EMT). This results in the loss of cell polarity and intercellular adhesion properties, enabling elastic morphological deformations of the cancer cells. These factors facilitate detachment of the cancer cells from the primary tumour and invasion through the basal membrane into the blood or lymphatic system – also known as intravasation (**Figure 1.4a**).⁶⁰ Once inside the circulatory system, tumour cells are subjected to several stresses such as haemodynamic forces, collisions with host cells and immunosurveillance. Only a fraction of these circulating tumour cells will thus survive to form metastases.⁶¹

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

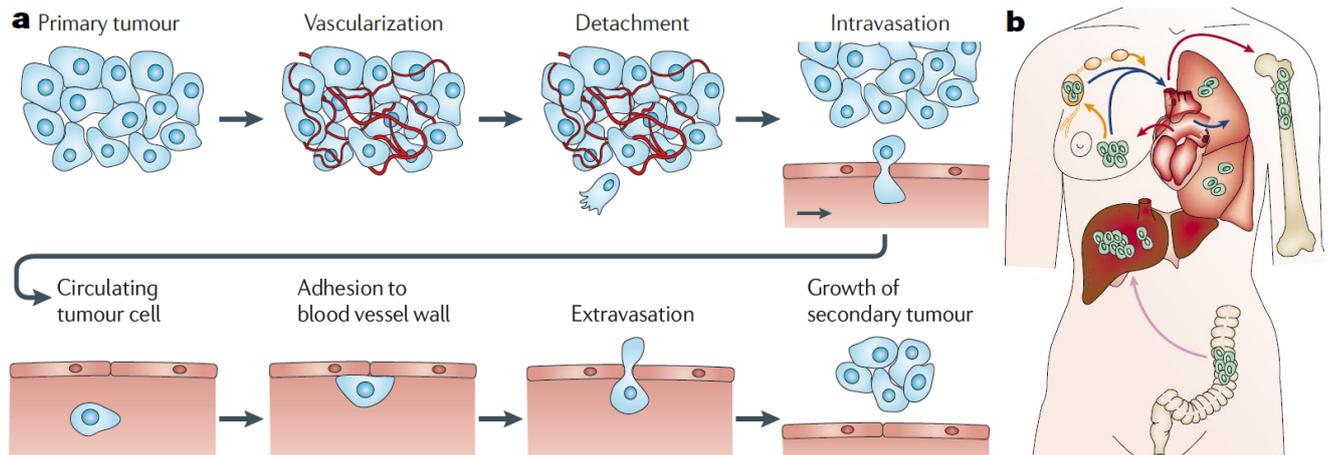


Figure 1.4: a) The metastatic process resulting in invasion and formation of secondary tumours b) Vascular and lymphatic flow patterns and the corresponding movement of cancer cells and tumour formation in favoured organs. Figures reproduced from a) Wirtz and b) Chambers et al.^{2, 60}

This survival requires adhesion to a blood vessel wall and extravasation from the circulation into the surrounding tissue which is largely dependent on the cancer cell and capillaries diameter.⁶⁰ After arrest, their ability to grow in secondary sites is governed by molecular interactions within the microenvironment of the new organ.

It is recognized that different types of cancer show an organ-specific pattern of metastasis, as exemplified in **Figure 1.4b**.² Breast cancer is frequently observed to metastasize in the lung, brain and bone of patients, while in contrast colorectal cancer preferentially spreads to the liver. These predispositions derive from a combination of vascular flow patterns and the theory of “seed and soil” or the propensity of a cancer cell (seed) to flourish in a specific secondary organ (soil). As an example, breast cancer cells leaving the primary tumour by blood vessels will flow firstly through the heart and then to the capillary bed of the lung (**Figure 1.4b**, blue arrows). Some cells may form metastases within the lung, while others may pass through the tissue and enter the arterial system (**Figure 1.4b**, red arrows), spreading to remote areas such as the brain and bones. Breast cancer cells have been known to invade the lymphatic system (**Figure 1.4b**, yellow arrows), taken to the draining lymph node and then entering the venous system. By contrast, colorectal cancer cells travel via the hepatic-portal circulatory system (**Figure 1.4b**, pink arrow) and are taken first to the capillaries of the liver.²

While still an emerging field with many unanswered questions, research into the mechanisms of this complex hallmark capability has gained traction over the past two decades. New technologies and research tools, coupled with refined experimental models and the identification of critical regulatory genes and interactions, has made significant progress possible in unravelling the important features of metastasis.

1.2.3 Emerging Hallmarks

The hallmark principles of cancer have been defined as capabilities that ensure cancer cell survival, proliferation and growth.¹⁵ The acquisition and activation of these functions, in different tumour types at different stages of tumorigenesis, is made possible by the two chief enabling characteristics outlined previously, namely: genome instability and inflammation within the tumour microenvironment as shown on the following page in **Figure 1.5**. In recent years, the addition of two other distinct attributes of cancer cells have been proposed.

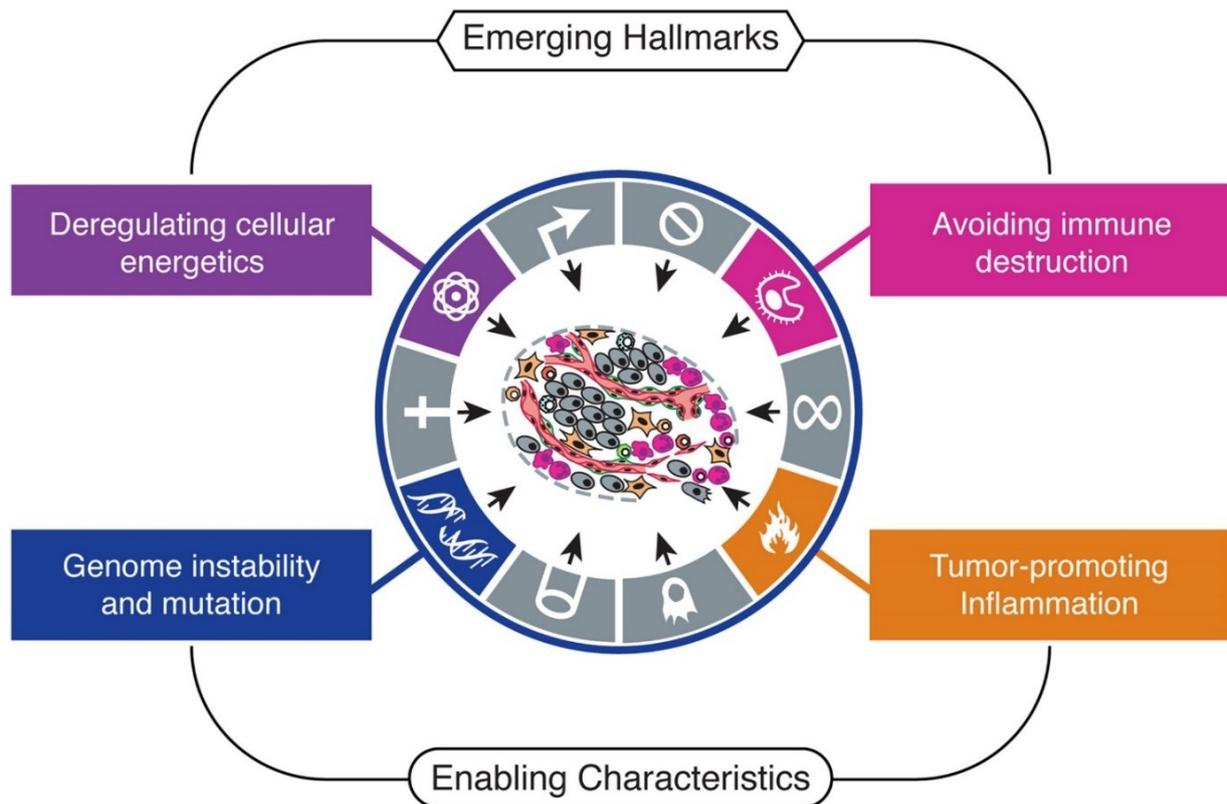


Figure 1.5: The two main enabling characteristics of genome instability and inflammation are shown above. Both cellular metabolism and immuno-evasion have also been recognised as hallmarks. Figure reproduced from Hanahan and Weinberg.¹⁵

These two features, illustrated in **Figure 1.5**, involve the reprogramming of cellular metabolism and energetics to sustain the uncontrolled and continuous cell growth associated with tumours, as well as the active ability of cancer cells to evade attack and elimination by the host's immune system. These capabilities have been deemed functionally important to tumour growth and progression in various forms of human cancer and will be discussed briefly.

1.2.3.1 Cellular Metabolism

Genetic alterations which drive abnormal cellular proliferation in tumours place a significant demand on nutrient utilisation and biosynthetic requirements for increased cell growth and division. Corresponding energy metabolism adjustments allow malignant cells to meet these requirements. Under aerobic conditions, normal cells rely primarily on mitochondrial oxidative phosphorylation to generate energy. In contrast, malignant tumours rely on glycolysis in the presence or absence of oxygen, a phenomenon first observed by Otto Warburg in 1930 and termed the "Warburg effect."⁶² A functional rationale for this glycolytic switch in cancer cells has been enigmatic in the past decades, owing to the reduced efficiency of generating adenosine 5'-triphosphate (ATP) in comparison to normal mitochondrial oxidative phosphorylation. It is now understood that the metabolism of cancer cells has adapted for this increased glycolysis, allowing for an increased uptake and incorporation of glycolytic intermediates into various biosynthetic pathways. This in turn facilitates the biosynthesis of macromolecules and organelles necessary for production of new cells.⁶² Mutations within signalling pathways implicated in cell growth have been shown to regulate these metabolic pathways and enable cancer cells to obtain and metabolise nutrients in a way conducive to cell division, rather than efficient ATP production.⁶² The reprogramming of energy metabolism has been highlighted as an emerging hallmark of tumorigenesis and key to its promotion within neoplastic tissue.¹⁵

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

1.2.3.2 Immune Evasion

The theorem of immune surveillance postulates that cells and tissues are continually being policed by the immune system, recognizing arising mutations and eradicating the vast majority of incipient neoplasms.⁶³ However, tumours that do survive and spread have done so by managing to avoid detection and evade the immunological eradication response. Recently, studies involving genetically engineered animal models and clinical epidemiology, have increased our understanding in this area and began to elucidate this phenomenon.⁶⁴ The physiological barriers which impede an antitumor response operate within the tumour microenvironment and make use of various components of the innate immune system. On such impediment, immunoediting, allows tumours to lay dormant through equilibrium and senescence before resurfacing. Others include inhibiting the secretions of regulatory T-cells, modifying production of immune suppressive mediators and immune deviation.⁶⁵ The immune evasion capability of tumours, and its significance in tumour initiation and progression, has established it as an emerging hallmark and enabling factor in the pathogenesis of cancer.¹⁵

1.3 Chemo- and Targeted Therapy of Cancer

Of all the obstacles faced and progress made in the field of medicine, few have been as arduous and hard-fought as the treatment and cure of cancer. Tremendous strides made in molecular cancer biology and genomics has expedited the development and transformation of cancer therapeutics from a low-budget, government-run research endeavour to a lucrative, multibillion-dollar industry. The intellectual foundations of this revolution can be traced back 110 years ago, when the founder of chemotherapy, Paul Ehrlich, received the Nobel Prize for Physiology or Medicine. His coining of the term “magic bullets”, for the treatment of human diseases, inspired generations of scientists and led to multiple ground-breaking studies, paving the road towards modern targeted disease therapy.⁶⁶ A chronicle of the history of chemo- and targeted therapy, focusing on clinical successes and remaining challenges and limitations within the field, will be discussed within this section. Finally, we will consider how the previously outlined hallmark principles of cancer have begun to shape the landscape of therapeutic development in oncology and will continue to do so in the future.

1.3.1 History

1.3.1.1 Chemotherapy

Prior to two major discoveries in the 1940s, cancer treatment remained largely the territory of the surgeon. The invention of the linear accelerator enabled widespread use of radiation therapy in the 1960s, but, as with surgery, failed to eradicate the spread and metastasis of tumours. The first four decades of the 20th century saw breakthroughs in screening methods and model development. In 1910, George Clowes successfully developed the first transplantable tumour system for rodents, allowing for standardisation and model systems for cancer drug testing.⁶⁷ Murray Shear later initiated the most organised program for cancer drug screening in 1935, leading to the testing of over 3000 compounds, including natural products.⁶⁷ The dawn of the modern era of chemotherapy can be traced back to two breakthroughs during World War II, the first of these being the discovery of nitrogen mustard as an efficacious cancer therapeutic by Louis Goodman and Alfred Gilman.⁶⁸

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

Autopsies of soldiers exposed to mustard gas in World War I revealed markedly depleted bone marrow and lymph nodes. This led the pair to experiments with nitrogen mustard on mice bearing transplanted lymphoid tumours, resulting in tumour regression. These results were successfully replicated in a patient with non-Hodgkin's lymphoma. The remission was short lived, and the disease progressed, but the principle of drug induced tumour regression was established.⁶⁸ The investigation into the effects of folic acid on patients with acute lymphoblastic leukaemia (ALL) by the pathologist Sydney Faber is regarded as the second major discovery. His collaboration with medicinal chemists in the synthesis of folate analogues led to the folate antagonist methotrexate, which acts by blocking the enzymes that require folic acid to synthesize DNA and promote cell division. Farber administered methotrexate to children with ALL in the late 1940s, it becoming the first drug to successfully induce remission in the disease.⁶⁷

In the early 1950s, several other antileukemic agents were developed which reached clinical trials and eventually induced long term remission in patients with ALL. This success resulted in legislation by the United States to initiate the National Cancer Chemotherapy Service Centre (NCCSC) at the National Cancer Institute (NCI) in 1955, known today as the Cancer Therapy Evaluation Program (CTEP).⁶⁸ This program continues to play a critical role in development of cancer therapeutics to this day. Combination therapy began in the 1960s for treatment of childhood ALL and was extended to lymphomas with startling results - complete remission rates jumped from near zero to over 80%, with some patients still not showing relapse after 40 years.⁶⁷ For the first time in cancer treatment the concept of cure became a reality. This advance was a scientific boon and had a remarkably permissive effect for the use of chemotherapy, resulting in the passing of the National Cancer Act of 1971 and President Nixon's declaration of the "war on cancer."

While there were successes in the period from 1970 to 1990, such as the discovery of cisplatin and improved screening methods, researchers continually encountered problems with the long-term toxicity of chemotherapeutic agents affecting nearly every organ. Less than 10% of new drugs that entered clinical trials attained FDA approval and mouse models for solid tumours were unreliable in predicting clinical success and cytotoxicity.⁶⁸ There was thus a clear need for a change in strategy in the fight against cancer.

1.3.1.2 The Targeted Therapy Revolution

While attempts to improve the stagnation of chemotherapeutics proceeded in the late 1980s, an understanding of cancer biology using molecular and genetic approaches was beginning to be uncovered by cell biologists. The discovery of wholly new signalling networks, responsible for the regulation of cell proliferation and survival, were found to be dramatically altered in cancer cells. This sparked an industrial revolution, particularly in small biotech start-ups, as scientists set out to target these molecular switches and defects including signalling molecules, growth factors, apoptosis modulators and angiogenesis promoters.⁶⁸ By the 1990s, the explosion of new drug targets ushered in the era of targeted therapy.

The success of finding inhibitors for these new targets was compounded by innovations in technology at the time and the emergence of combinatorial chemistry - furnishing thousands of novel scaffolds for high throughput screening (HTS). Lead compounds were identified and optimized for enhanced bioavailability and selectivity. Throughout this period, it became clear that promising drug candidates should exhibit certain favourable characteristics.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

These include metabolic stability, good oral bioavailability and low toxicity profile at effective doses. Adoption of this approach led to marked successes within the targeted therapy industry in the ensuing decades.

A landmark event in the targeted therapy revolution was the development of the natural product derived imatinib mesylate, a structurally simple compound bearing all the desirable features of an ideal targeted therapeutic.⁶⁹ Synthesized by chemists at Novartis, imatinib is used to treat CML and ALL by inhibiting the kinase BCR-ABL, a fusion protein produced by the Philadelphia chromosomal translocation. Imatinib has proven remarkably successful, with over 90% of chronic-phase patients achieving complete haematological remission.⁶⁹ It received FDA approval in 2001 and has merited a position on the World Health Organisation's List of Essential Medicines.⁷⁰ In the acute-phase of CML, treatment with imatinib produces brief remission, which is followed by the rapid emergence of resistant cells that arise from mutations in the catalytic domain of the kinase. These drug resistant mutations were also found to be present in patients before treatment, providing valuable insights into the existence of inherent mutations and drug-resistant mutant subclones in cancer.⁶⁸

Another controversial class of compounds, which inhibit the epidermal growth factor receptor, has secured numerous FDA approvals over the last 15 years. Used in the treatment of NSCLC and breast cancer, these inhibitors have been plagued by the occurrence of mutations within the kinase active site which they target, rendering the agents largely ineffective.⁷¹ The discovery and efforts to overcome the mutations that nullify EGFR inhibitors has represented an important advancement in targeted therapy and forms the backbone of this body of research. An overview of this therapeutic target and its targeted therapy will be discussed in detail in the next chapter.

1.3.2 Modern Targeted Therapy

1.3.2.1 Industry Overview

In the last two decades, oncology has experienced a tidal wave of innovation that has profoundly changed how we diagnose and treat cancer. In 2003, the worldwide expenditure for oncology-related and supportive care medicines was \$36 billion.⁶⁸ This number increased to \$107 billion in 2015, a 300% market gain in 12 years, and is expected to exceed \$150 billion by 2020.⁷² Oncology is a key driver of growth in the pharmaceutical industry and is projected to account for approximately 30% of the industries' product pipeline and 25% of the total revenue in 2020.⁷²

Spanning the years from 1949 to 2015, a total of 150 anticancer drugs received FDA approval, as depicted in the graph in **Figure 1.6** on the following page.⁷³ This includes a total of 61 cytotoxic and 89 target-based drugs used to treat 33 different cancer types. A clear decline in cytotoxic and a rise in target-based drug approval and development can be seen from 1990 onwards, with over 90% of all target-based therapeutics receiving approval in the last 20 years. In 2015, the WHO Model List of Essential Medicines added a total of 16 new anticancer drugs, targeting 26 different cancer types.⁷⁴ This was the largest ever addition to any therapeutic area since the lists inception in 1997, with first-time inclusions of 3 target-based agents, namely: imatinib, rituximab and trastuzumab.⁷⁰ This amount was further supplemented by the addition of the two oral inhibitors, dasatinib and nilotinib, bringing the total number of essential medicines for cancer to 48.⁷⁰

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

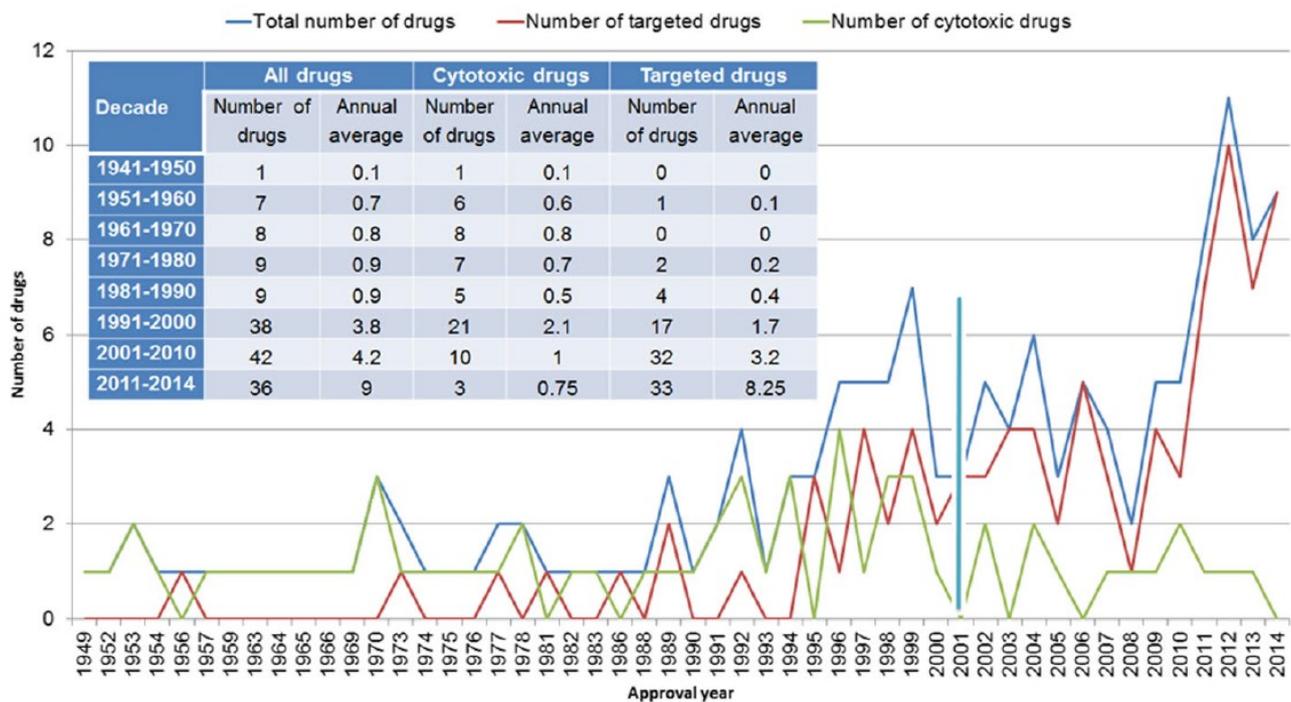


Figure 1.6: Anticancer drugs that have received FDA approval from 1949 to 2014. Graph reproduced from Sun et al.⁷³

1.3.2.2 Challenges and Limitations

The establishment of targeted therapy in the treatment of human cancer is the remuneration of over three decades worth of research into the mechanisms surrounding the pathogenesis of cancer. This achievement has drawn intellectual momentum from the initial discoveries of Paul Ehrlich and his ideal of “aiming precisely.” Using a rational, targeted strategy remains the paradigm for modern cancer research and drug discovery today. The transition from cytotoxic drugs to targeted therapies constitutes a significant leap and we now have more and better drugs than ever before. However, cancer remains a multi-dimensional group of diseases and its complexity is revealed through resistance and off-target issues. This raises the question of requirement for such a “magic bullet” in this genetically complicated scenario.

Most cancers possess numerous driver mutations and increasing evidence suggests that the use of drugs targeting multiple oncogenic functions offers a promising approach over the current stance of mono-specific drugs being used to overcome multifactorial diseases. However, this increase in the number of pathways targeted by a single compound raises the risk of adverse and off-target effects. Finding a balance between target specificity and efficacy remains a major challenge in drug development. The mutability of cancer cells and the onset of resistance to these compounds remains an unpredictable and devastating obstacle. A major scientific challenge of the future will be to find ways to predict and exploit these occurrences. It is now clear that understanding the molecular profile of human tumours is fundamental in the effective design and use of cancer drugs. Patient selection, and the use of molecular and genomic assays to identify subclones present within these patients, provides an enormous advantage in clinical design and increases the likelihood of successful treatment response. This also avoids the undesirable expense and toxicity of ineffective treatment, leading to cost-effective drug development. The hallmark principles of cancer, and our growing knowledge of their inner workings, are fundamental in cancer drug development at present and will continue to be in the future of patient-tailored molecular therapy.

1.3.2.3 Targeting the Hallmarks

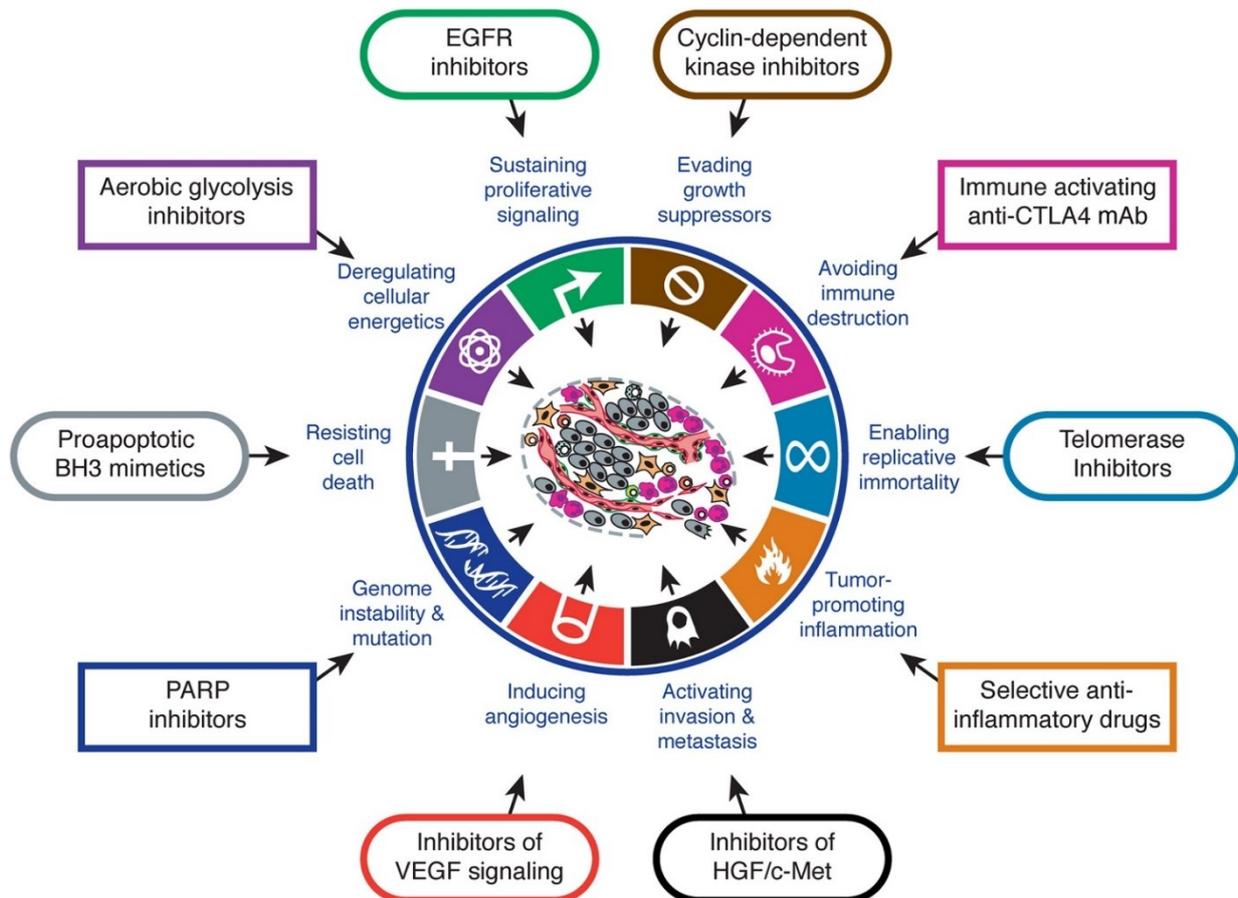


Figure 1.7: Therapeutic targeting of each of the hallmark capabilities by various inhibitors and agents. Figure reproduced from Hanahan and Weinberg.¹⁵

While it is impossible to enumerate the myriad of targeted therapies that are in development or being introduced to patients globally, this expanding armamentarium can be categorised according to their respective effects on certain hallmark capabilities of cancer cells, as exemplified above in **Figure 1.7**.¹⁵ The observed efficacy of these drugs in each category validates the importance of that capability in tumour initiation and progression, and thus its targeting.

Notably, most hallmark-targeting drugs have been developed to focus on one specific target involved in the enabling of a certain capability. While this is advantageous with regards to specificity and off-target toxicity, clinical responses have been temporary and inevitably end in relapse. Increasing experimental evidence suggests that inhibiting one key pathway in a tumour is not enough to completely stop a hallmark capability, allowing a subpopulation of cancer cells to survive with partial function and eventually to adapt and overcome the selective pressure imposed by the therapeutic.¹⁵ This re-establishment of a functional hallmark capability is known as adaptive resistance and may occur through mutations, epigenetic events or tumour microenvironment remodelling. As the number of parallel pathways promoting a hallmark must be finite, it is therefore theoretically possible to inhibit all these pathways therapeutically and avert adaptive resistance.

Another form of acquired drug resistance exhibited by cancer cells is the ability to reduce their reliance on a specific targeted hallmark capability and shift their dependence to another. This has been unexpectedly demonstrated in recent responses to angiogenesis inhibitors.⁷⁵

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

Clinical evaluation of these inhibitors exhibited successful suppression of angiogenic capabilities, only to be followed by a ramped-up activity in invasiveness and metastasis which allowed for access to pre-existing vasculature in nearby tissues. The applicability of this resistance is yet to be established in other cancers, but the concept of selective, mechanism-guided combination targeting of multiple hallmark capabilities could result in more efficacious and durable targeted therapeutics.

The drug classes listed in **Figure 1.7** are illustrative examples of a burgeoning pool of candidate and approved drugs with different molecular targets and mechanisms of action. These include targeting apoptosis,⁷⁶ inflammation,⁵⁷ immunotherapy⁷⁷ or epigenetics⁷⁸ amongst a host of others. It has been recognized that targeting the crucial signalling proteins that regulate hallmark and enabling characteristics of cancer is the ultimate target for molecular therapy.⁶⁶ The body of research presented in this thesis focusses on the EGFR signalling pathway, its implication in cancer and the value of inhibiting the kinase domain of this receptor in the treatment of cancer. The following chapter will explore structural and functional features of the receptor, its oncogenic activation and its history as a therapeutic target in the development of tumour growth inhibiting drugs.

1.4 References

1. C. J. Sherr, *Science*, 1996, **274**, 1672-1677.
2. A. F. Chambers, A. C. Groom and I. C. MacDonald, *Nature Reviews Cancer*, 2002, **2**, 563-572.
3. *World Health Statistics 2018: Monitoring Health for the SDGs - Sustainable Development Goals*, World Health Organization, 2018.
4. *Global Status Report on Noncommunicable Diseases 2014*, World Health Organization, 2014.
5. M. H. Forouzanfar, A. Afshin, L. T. Alexander, H. R. Anderson, Z. A. Bhutta, S. Biryukov, M. Brauer, R. Burnett, K. Cercy, F. J. Charlson, A. J. Cohen, L. Dandona, K. Estep, A. J. Ferrari, J. J. Frostad, N. Fullman, P. W. Gething, W. W. Godwin, M. Griswold, S. I. Hay, Y. Kinfu, H. H. Kyu, H. J. Larson, X. Liang, S. S. Lim, P. Y. Liu, A. D. Lopez, R. Lozano, L. Marczak, G. A. Mensah, A. H. Mokdad, M. Moradi-Lakeh, M. Naghavi, B. Neal, M. B. Reitsma, G. A. Roth, J. A. Salomon, P. J. Sur, T. Vos, J. A. Wagner, H. Wang, Y. Zhao, M. Zhou, G. M. Aasvang, A. A. Abajobir, K. H. Abate, C. Abbafati, K. M. Abbas, F. Abd-Allah, A. M. Abdulle, S. F. Abera, B. Abraham, L. J. Abu-Raddad, G. Y. Abyu, A. O. Adebisi, I. A. Adedeji, Z. Ademi, A. K. Adou, J. C. Adsuar, E. E. Agardh, A. Agarwal, A. Agrawal, A. A. Kiadaliri, O. N. Ajala, T. F. Akinyemiju, Z. Al-Aly, K. Alam, N. K. M. Alam, S. F. Aldhahri, R. W. Aldridge, Z. A. Alemu, R. Ali, A. a. Alkerwi, F. Alla, P. Allebeck, U. Alsharif, K. A. Altirkawi, E. A. Martin, N. Alvis-Guzman, A. T. Amare, A. Amberbir, A. K. Amegah, H. Amini, W. Ammar, S. M. Amrock, H. H. Andersen, B. O. Anderson, C. A. T. Antonio, P. Anwari, J. Ärnlöv, A. Artaman, H. Asayesh, R. J. Asghar, R. Assadi, S. Atique, E. F. G. A. Avokpaho, A. Awasthi, B. P. A. Quintanilla, P. Azzopardi, U. Bacha, A. Badawi, M. C. Bahit, K. Balakrishnan, A. Barac, R. M. Barber, S. L. Barker-Collo, T. Bärnighausen, S. Barquera, L. Barregard, L. H. Barrero, S. Basu, C. Batis, S. Bazargan-Hejazi, J. Beardsley, N. Bedi, E. Beghi, B. Bell, M. L. Bell, A. K. Bello, D. A. Bennett, I. M. Bensenor, A. Berhane, E. Bernabé, B. D. Betsu, A. S. Beyene, N. Bhala, A. Bhansali, S. Bhatt, S. Biadgilign, B. Bikbov, D. Bisanzio, E. Bjertness, J. D. Blore, R. Borschmann, S. Boufous, R. R. A. Bourne, M. Brainin, A. Brazinova, N. J. K. Breitborde, H. Brenner, D. M. Broday, T. S. Brugha, B. Brunekreef, Z. A. Butt, L. E. Cahill, B. Calabria, I. R. Campos-Nonato, R. Cárdenas, D. O. Carpenter, J. J. Carrero, D. C. Casey, C. A. Castañeda-Orjuela, J. C. Rivas, R. E. Castro, F. Catalá-López, J.-C. Chang, P. P.-C. Chiang, M. Chibalabala, O. Chimed-Ochir, V. H. Chisumpa, A. A. Chitheer, J.-Y. J. Choi, H. Christensen, D. J. Christopher, L. G. Ciobanu, M. M. Coates, S. M. Colquhoun, A. G. C. Manzano, L. T. Cooper, K. Cooperrider, L. Cornaby, M. Cortinovis, J. A. Crump, L. Cuevas-Nasu, A. Damasceno, R. Dandona, S. C. Darby, P. I. Dargan, J. das Neves, A. C. Davis, K. Davletov, E. F. de Castro, V. De la Cruz-Góngora, D. De Leo, L. Degenhardt, L. C. Del Gobbo, B. del Pozo-Cruz, R. P. Dellavalle, A. Deribew, D. C. D. Jarlais, S. D. Dharmaratne, P. K. Dhillon, C. Diaz-Torné, D. Dicker, E. L. Ding, E. R. Dorsey, K. E. Doyle, T. R. Driscoll, L. Duan, M. Dubey, B. B. Duncan, I. Elyazar, A. Y. Endries, S. P. Ermakov, H. E. Erskine, B. Eshrati, A. Esteghamati, S. Fahimi, E. J. A. Faraon, T. A. Farid, C. S. e. S. Farinha, A. Faro, M. S. Farvid, F. Farzadfar, V. L. Feigin, S.-M. Fereshtehnejad, J. G. Fernandes, F. Fischer, J. R. A. Fitchett, T. Fleming, N. Foigt, K. Foreman, F. G. R. Fowkes, R. C. Franklin, T.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

Fürst, N. D. Futran, E. Gakidou, A. L. Garcia-Basteiro, T. T. Gebrehiwot, A. T. Gebremedhin, J. M. Geleijnse, B. D. Gessner, A. Z. Giref, M. Giroud, M. D. Gishu, G. Giussani, S. Goenka, M. C. Gomez-Cabrera, H. Gomez-Dantes, P. Gona, A. Goodridge, S. V. Gopalani, C. C. Gotay, A. Goto, H. N. Gouda, H. C. Gughani, F. Guillemin, Y. Guo, R. Gupta, R. Gupta, R. A. Gutiérrez, J. A. Haagsma, N. Hafezi-Nejad, D. Haile, G. B. Hailu, Y. A. Halasa, R. R. Hamadeh, S. Hamidi, A. J. Handal, G. J. Hankey, Y. Hao, H. L. Harb, S. Harikrishnan, J. M. Haro, M. S. Hassanvand, T. A. Hassen, R. Havmoeller, I. B. Heredia-Pi, N. F. Hernández-Llanes, P. Heydarpour, H. W. Hoek, H. J. Hoffman, M. Horino, N. Horita, H. D. Hosgood, D. G. Hoy, M. Hsairi, A. S. Htet, G. Hu, J. J. Huang, A. Hussein, S. J. Hutchings, I. Huybrechts, K. M. Iburg, B. T. Idrisov, B. V. Ileanu, M. Inoue, T. A. Jacobs, K. H. Jacobsen, N. Jahanmehr, M. B. Jakovljevic, H. A. F. M. Jansen, S. K. Jassal, M. Javanbakht, S. P. Jayaraman, A. U. Jayatilleke, S. H. Jee, P. Jeemon, V. Jha, Y. Jiang, T. Jibat, Y. Jin, C. O. Johnson, J. B. Jonas, Z. Kabir, Y. Kalkonde, R. Kamal, H. Kan, A. Karch, C. K. Karema, C. Karimkhani, A. Kasaeian, A. Kaul, N. Kawakami, D. S. Kazi, P. N. Keiyoro, L. Kemmer, A. H. Kemp, A. P. Kengne, A. Keren, C. N. Kesavachandran, Y. S. Khader, A. R. Khan, E. A. Khan, G. Khan, Y.-H. Khang, S. Khatibzadeh, S. Khera, T. A. M. Khoja, J. Khubchandani, C. Kieling, C.-i. Kim, D. Kim, R. W. Kimokoti, N. Kissoon, M. Kivipelto, L. D. Knibbs, Y. Kokubo, J. A. Kopec, P. A. Koul, A. Koyanagi, M. Kravchenko, H. Kromhout, H. Krueger, T. Ku, B. K. Defo, R. S. Kuchenbecker, B. K. Bicer, E. J. Kuipers, G. A. Kumar, G. F. Kwan, D. K. Lal, R. Laloo, T. Lallukka, Q. Lan, A. Larsson, A. A. Latif, A. E. B. Lawrynowicz, J. L. Leasher, J. Leigh, J. Leung, M. Levi, X. Li, Y. Li, J. Liang, S. Liu, B. K. Lloyd, G. Logroscino, P. A. Lotufo, R. Lunevicius, M. MacIntyre, M. Mahdavi, M. Majdan, A. Majeed, R. Malekzadeh, D. C. Malta, W. A. A. Manamo, C. C. Mapoma, W. Marcenes, R. V. Martin, J. Martinez-Raga, F. Masiye, K. Matsushita, R. Matzopoulos, B. M. Mayosi, J. J. McGrath, M. McKee, P. A. Meaney, C. Medina, A. Mehari, F. Mejia-Rodriguez, A. B. Mekonnen, Y. A. Melaku, Z. A. Memish, W. Mendoza, G. B. M. Mensink, A. Meretoja, T. J. Meretoja, Y. M. Mesfin, F. A. Mhimbira, A. Millea, T. R. Miller, E. J. Mills, M. Mirarefin, A. Misganaw, C. N. Mock, A. Mohammadi, S. Mohammed, G. L. D. Mola, L. Monasta, J. C. M. Hernandez, M. Montico, L. Morawska, R. Mori, D. Mozaffarian, U. O. Mueller, E. Mullany, J. E. Mumford, G. V. S. Murthy, J. B. Nachega, A. Naheed, V. Nangia, N. Nassiri, J. N. Newton, M. Ng, Q. L. Nguyen, M. I. Nisar, P. M. N. Pete, O. F. Norheim, R. E. Norman, B. Norrving, L. Nyakarahuka, C. M. Obermeyer, F. A. Ogbo, I.-H. Oh, O. Oladimeji, P. R. Olivares, H. Olsen, B. O. Olusanya, J. O. Olusanya, J. N. Opio, E. Oren, R. Orozco, A. Ortiz, E. Ota, M. Pa, A. Pana, E.-K. Park, C. D. Parry, M. Parsaeian, T. Patel, A. J. P. Caicedo, S. T. Patil, S. B. Patten, G. C. Patton, N. Pearce, D. M. Pereira, N. Perico, K. Pesudovs, M. Petzold, M. R. Phillips, F. B. Piel, J. D. Pillay, D. Plass, S. Polinder, C. D. Pond, C. A. Pope, D. Pope, S. Popova, R. G. Poulton, F. Pourmalek, N. M. Prasad, M. Qorbani, R. H. S. Rabiee, A. Radfar, A. Rafay, V. Rahimi-Movaghar, M. Rahman, M. H. U. Rahman, S. U. Rahman, R. K. Rai, S. Rajsic, M. Raju, U. Ram, S. M. Rana, K. Ranganathan, P. Rao, C. A. R. García, A. H. Refaat, C. D. Rehm, J. Rehm, N. Reinig, G. Remuzzi, S. Resnikoff, A. L. Ribeiro, J. A. Rivera, H. S. Roba, A. Rodriguez, S. Rodriguez-Ramirez, D. Rojas-Rueda, Y. Roman, L. Ronfani, G. Roshandel, D. Rothenbacher, A. Roy, M. M. Saleh, J. R. Sanabria, L. Sanchez-Riera, M. D. Sanchez-Niño, T. G. Sánchez-Pimenta, L. Sandar, D. F. Santomauro, I. S. Santos, R. Sarmiento-Suarez, B. Sartorius, M. Satpathy, M. Savic, M. Sawhney, J. Schmidhuber, M. I. Schmidt, I. J. C. Schneider, B. Schöttker, A. E. Schutte, D. C. Schwebel, J. G. Scott, S. Seedat, S. G. Sepanlou, E. E. Servan-Mori, G. Shaddick, A. Shaheen, S. Shahrzad, M. A. Shaikh, T. S. Levy, R. Sharma, J. She, S. Sheikhabaehi, J. Shen, K. N. Sheth, P. Shi, K. Shibuya, M. Shigematsu, M.-J. Shin, R. Shiri, K. Shishani, I. Shiue, M. G. Shrima, I. D. Sigfusdottir, D. A. S. Silva, D. G. A. Silveira, J. I. Silverberg, E. P. Simard, S. Sindi, A. Singh, J. A. Singh, P. K. Singh, E. L. Slepak, M. Soljak, S. Soneji, R. J. D. Sorensen, L. A. Sposato, C. T. Sreeramareddy, V. Stathopoulou, N. Steckling, N. Steel, D. J. Stein, M. B. Stein, H. Stöckl, S. Stranges, K. Stroumpoulis, B. F. Sunguya, S. Swaminathan, B. L. Sykes, C. E. I. Szoeki, R. Tabarés-Seisdedos, K. Takahashi, R. T. Talongwa, N. Tandon, D. Tanne, M. Tavakkoli, B. W. Taye, H. R. Taylor, B. A. Tedla, W. M. Tefera, T. K. Tegegne, D. Y. Tekle, A. S. Terkawi, J. S. Thakur, B. A. Thomas, M. L. Thomas, A. J. Thomson, A. L. Thorne-Lyman, A. G. Thrift, G. D. Thurston, T. Tillmann, R. Tobe-Gai, M. Tobollik, R. Topor-Madry, F. Topouzis, J. A. Towbin, B. X. Tran, Z. T. Dimbuene, N. Tsilimparis, A. K. Tura, E. M. Tuzcu, S. Tyrovolas, K. N. Ukwaja, E. A. Undurraga, C. J. Uneke, O. A. Uthman, A. van Donkelaar, J. van Os, Y. Y. Varakin, T. Vasankari, J. L. Veerman, N. Venketasubramanian, F. S. Violante, S. E. Vollset, G. R. Wagner, S. G. Waller, J. L. Wang, L. Wang, Y. Wang, S. Weichenthal, E. Weiderpass, R. G. Weintraub, A. Werdecker, R. Westerman, H. A. Whiteford, T. Wijeratne, C. S. Wiysonge, C. D. A. Wolfe, S. Won, A. D. Woolf, M. Wubshet, D. Xavier, G. Xu, A. K. Yadav, B. Yakob, A. Z. Yalew, Y. Yano, M. Yaseri, P. Ye, P. Yip, N. Yonemoto, S.-J. Yoon, M. Z. Younis, C. Yu, Z. Zaidi, M. E. S. Zaki, J. Zhu, B. Zipkin, S. Zodpey, L. J. Zuhlke and C. J. L. Murray, *The Lancet*, 2016, **388**, 1659-1724.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

6. *Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000-2016*, World Health Organization, 2018.
7. *World Cancer Report 2014*, International Agency for Research on Cancer, 2014.
8. J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman and F. Bray, *International Journal of Cancer*, 2015, **136**, E359-E386.
9. M. G. Valsecchi and E. Steliarova-Foucher, *The Lancet Oncology*, 2008, **9**, 159-167.
10. K. A. Fleming, M. Naidoo, M. Wilson, J. Flanigan, S. Horton, M. Kuti, L. M. Looi, C. Price, K. Ru, A. Ghafur, J. Wang and N. Lago, *American Journal of Clinical Pathology*, 2017, **147**, 15-32.
11. M. Plummer, C. de Martel, J. Vignat, J. Ferlay, F. Bray and S. Franceschi, *The Lancet Global Health*, 2016, **4**, E609-E616.
12. *New Perspectives on Global Health Spending for Universal Health Coverage*, World Health Organization, 2017.
13. A. Pearce, L. Sharp, P. Hanly, A. Barchuk, F. Bray, M. de Camargo Cancela, P. Gupta, F. Meheus, Y.-L. Qiao, F. Sitas, S.-M. Wang and I. Soerjomataram, *Cancer Epidemiology*, 2018, **53**, 27-34.
14. *The Economics of Cancer Prevention and Control Data Digest*, Union for International Cancer Control, 2014.
15. D. Hanahan and Robert A. Weinberg, *Cell*, 2011, **144**, 646-674.
16. K. A. Hoadley, C. Yau, T. Hinoue, D. M. Wolf, A. J. Lazar, E. Drill, R. Shen, A. M. Taylor, A. D. Cherniack, V. Thorsson, R. Akbani, R. Bowlby, C. K. Wong, M. Wiznerowicz, F. Sanchez-Vega, A. G. Robertson, B. G. Schneider, M. S. Lawrence, H. Noushmehr, T. M. Malta, S. J. Caesar-Johnson, J. A. Demchok, I. Felau, M. Kasapi, M. L. Ferguson, C. M. Hutter, H. J. Sofia, R. Tarnuzzer, Z. Wang, L. Yang, J. C. Zenklusen, J. Zhang, S. Chudamani, J. Liu, L. Lolla, R. Naresh, T. Pihl, Q. Sun, Y. Wan, Y. Wu, J. Cho, T. DeFreitas, S. Frazer, N. Gehlenborg, G. Getz, D. I. Heiman, J. Kim, M. S. Lawrence, P. Lin, S. Meier, M. S. Noble, G. Saksena, D. Voet, H. Zhang, B. Bernard, N. Chambwe, V. Dhankani, T. Knijnenburg, R. Kramer, K. Leinonen, Y. Liu, M. Miller, S. Reynolds, I. Shmulevich, V. Thorsson, W. Zhang, R. Akbani, B. M. Broom, A. M. Hegde, Z. Ju, R. S. Kanchi, A. Korkut, J. Li, H. Liang, S. Ling, W. Liu, Y. Lu, G. B. Mills, K.-S. Ng, A. Rao, M. Ryan, J. Wang, J. N. Weinstein, J. Zhang, A. Abeshouse, J. Armenia, D. Chakravarty, W. K. Chatila, I. de Bruijn, J. Gao, B. E. Gross, Z. J. Heins, R. Kundra, K. La, M. Ladanyi, A. Luna, M. G. Nissan, A. Ochoa, S. M. Phillips, E. Reznik, F. Sanchez-Vega, C. Sander, N. Schultz, R. Sheridan, S. O. Sumer, Y. Sun, B. S. Taylor, J. Wang, H. Zhang, P. Anur, M. Peto, P. Spellman, C. Benz, J. M. Stuart, C. K. Wong, C. Yau, D. N. Hayes, J. S. Parker, M. D. Wilkerson, A. Ally, M. Balasundaram, R. Bowlby, D. Brooks, R. Carlsen, E. Chuah, N. Dhalla, R. Holt, S. J. M. Jones, K. Kasaian, D. Lee, Y. Ma, M. A. Marra, M. Mayo, R. A. Moore, A. J. Mungall, K. Mungall, A. G. Robertson, S. Sadeghi, J. E. Schein, P. Sipahimalani, A. Tam, N. Thiessen, K. Tse, T. Wong, A. C. Berger, R. Beroukhi, A. D. Cherniack, C. Cibulskis, S. B. Gabriel, G. F. Gao, G. Ha, M. Meyerson, S. E. Schumacher, J. Shih, M. H. Kucherlapati, R. S. Kucherlapati, S. Baylin, L. Cope, L. Danilova, M. S. Bootwalla, P. H. Lai, D. T. Maglinte, D. J. Van Den Berg, D. J. Weisenberger, J. T. Auman, S. Balu, T. Bodenheimer, C. Fan, K. A. Hoadley, A. P. Hoyle, S. R. Jefferys, C. D. Jones, S. Meng, P. A. Mieczkowski, L. E. Mose, A. H. Perou, C. M. Perou, J. Roach, Y. Shi, J. V. Simons, T. Skelly, M. G. Soloway, D. Tan, U. Veluvolu, H. Fan, T. Hinoue, P. W. Laird, H. Shen, W. Zhou, M. Bellair, K. Chang, K. Covington, C. J. Creighton, H. Dinh, H. Doddapaneni, L. A. Donehower, J. Drummond, R. A. Gibbs, R. Glenn, W. Hale, Y. Han, J. Hu, V. Korchina, S. Lee, L. Lewis, W. Li, X. Liu, M. Morgan, D. Morton, D. Muzny, J. Santibanez, M. Sheth, E. Shinbrot, L. Wang, M. Wang, D. A. Wheeler, L. Xi, F. Zhao, J. Hess, E. L. Appelbaum, M. Bailey, M. G. Cordes, L. Ding, C. C. Fronick, L. A. Fulton, R. S. Fulton, C. Kandoth, E. R. Mardis, M. D. McLellan, C. A. Miller, H. K. Schmidt, R. K. Wilson, D. Crain, E. Curley, J. Gardner, K. Lau, D. Mallery, S. Morris, J. Paulauskis, R. Penny, C. Shelton, T. Shelton, M. Sherman, E. Thompson, P. Yena, J. Bowen, J. M. Gastier-Foster, M. Gerken, K. M. Leraas, T. M. Lichtenberg, N. C. Ramirez, L. Wise, E. Zmuda, N. Corcoran, T. Costello, C. Hovens, A. L. Carvalho, A. C. de Carvalho, J. H. Fregnani, A. Longatto-Filho, R. M. Reis, C. Scapulatempo-Neto, H. C. S. Silveira, D. O. Vidal, A. Burnette, J. Eschbacher, B. Hermes, A. Noss, R. Singh, M. L. Anderson, P. D. Castro, M. Ittmann, D. Huntsman, B. Kohl, X. Le, R. Thorp, C. Andry, E. R. Duffy, V. Lyadov, O. Paklina, G. Setdikova, A. Shabunin, M. Tavobilov, C. McPherson, R. Warnick, R. Berkowitz, D. Cramer, C. Feltmate, N. Horowitz, A. Kibel, M. Muto, C. P. Raut, A. Malykh, J. S. Barnholtz-Sloan, W. Barrett, K. Devine, J. Fulop, Q. T. Ostrom, K. Shimmel, Y. Wolinsky, A. E. Sloan, A. De Rose, F. Giuliante, M. Goodman, B. Y. Karlan, C. H. Hagedorn, J. Eckman, J. Harr, J. Myers, K. Tucker, L. A. Zach, B. Deyarmin, H. Hu, L. Kvecher, C. Larson, R. J. Mural, S. Somiari, A. Vicha, T. Zelinka, J. Bennett, M. Iacocca, B. Rabeno, P. Swanson, M. Latour, L. Lacombe, B. Têtu, A. Bergeron, M. McGraw, S. M. Staugaitis, J. Chabot, H. Hibshoosh, A. Sepulveda, T. Su, T. Wang, O. Potapova, O. Voronina, L. Desjardins, O. Mariani, S. Roman-Roman, X. Sastre, M.-H. Stern, F. Cheng, S. Signoretti, A. Berchuck, D. Bigner, E. Lipp, J. Marks, S. McCall, R. McLendon, A. Secord, A. Sharp, M. Behera, D. J. Brat, A. Chen, K. Delman, S. Force, F. Khuri, K.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

- Magliocca, S. Maithel, J. J. Olson, T. Owonikoko, A. Pickens, S. Ramalingam, D. M. Shin, G. Sica, E. G. Van Meir, H. Zhang, W. Eijckenboom, A. Gillis, E. Korpershoek, L. Looijenga, W. Oosterhuis, H. Stoop, K. E. van Kessel, E. C. Zwarthoff, C. Calatuzzolo, L. Cuppini, S. Cuzzubbo, F. DiMeco, G. Finocchiaro, L. Mattei, A. Perin, B. Pollo, C. Chen, J. Houck, P. Lohavanichbutr, A. Hartmann, C. Stoehr, R. Stoehr, H. Taubert, S. Wach, B. Wullich, W. Kycler, D. Murawa, M. Wiznerowicz, K. Chung, W. J. Edenfield, J. Martin, E. Baudin, G. Bubley, R. Bueno, A. De Rienzo, W. G. Richards, S. Kalkanis, T. Mikkelsen, H. Noushmehr, L. Scarpace, N. Girard, M. Aymerich, E. Campo, E. Giné, A. L. Guillermo, N. Van Bang, P. T. Hanh, B. D. Phu, Y. Tang, H. Colman, K. Evason, P. R. Dottino, J. A. Martignetti, H. Gabra, H. Juhl, T. Akeredolu, S. Stepa, D. Hoon, K. Ahn, K. J. Kang, F. Beuschlein, A. Breggia, M. Birrer, D. Bell, M. Borad, A. H. Bryce, E. Castle, V. Chandan, J. Cheville, J. A. Copland, M. Farnell, T. Flotte, N. Giama, T. Ho, M. Kendrick, J.-P. Kocher, K. Kopp, C. Moser, D. Nagorney, D. O'Brien, B. P. O'Neill, T. Patel, G. Petersen, F. Que, M. Rivera, L. Roberts, R. Smallridge, T. Smyrk, M. Stanton, R. H. Thompson, M. Torbenson, J. D. Yang, L. Zhang, F. Brimo, J. A. Ajani, A. M. A. Gonzalez, C. Behrens, o. Bondaruk, R. Broaddus, B. Czerniak, B. Esmaeli, J. Fujimoto, J. Gershenwald, C. Guo, A. J. Lazar, C. Logothetis, F. Meric-Bernstam, C. Moran, L. Ramondetta, D. Rice, A. Sood, P. Tamboli, T. Thompson, P. Troncso, A. Tsao, I. Wistuba, C. Carter, L. Haydu, P. Hersey, V. Jakrot, H. Kakavand, R. Kefford, K. Lee, G. Long, G. Mann, M. Quinn, R. Saw, R. Scolyer, K. Shannon, A. Spillane, J. Stretch, M. Synott, J. Thompson, J. Wilmott, H. Al-Ahmadie, T. A. Chan, R. Ghossein, A. Gopalan, D. A. Levine, V. Reuter, S. Singer, B. Singh, N. V. Tien, T. Broudy, C. Mirsaidi, P. Nair, P. Drwiega, J. Miller, J. Smith, H. Zaren, J.-W. Park, N. P. Hung, E. Kebebew, W. M. Linehan, A. R. Metwalli, K. Pacak, P. A. Pinto, M. Schiffman, L. S. Schmidt, C. D. Vocke, N. Wentzensen, R. Worrell, H. Yang, M. Moncrieff, C. Goparaju, J. Melamed, H. Pass, N. Botnariuc, I. Caraman, M. Cernat, I. Chemencedji, A. Clipca, S. Doruc, G. Gorincioi, S. Mura, M. Pirtac, I. Stancul, D. Tcaciuc, M. Albert, I. Alexopoulou, A. Arnaout, J. Bartlett, J. Engel, S. Gilbert, J. Parfitt, H. Sekhon, G. Thomas, D. M. Rassl, R. C. Rintoul, C. Bifulco, R. Tamakawa, W. Urba, N. Hayward, H. Timmers, A. Antenucci, F. Facciolo, G. Grazi, M. Marino, R. Merola, R. de Krijger, A.-P. Gimenez-Roqueplo, A. Piché, S. Chevalier, G. McKercher, K. Birsoy, G. Barnett, C. Brewer, C. Farver, T. Naska, N. A. Pennell, D. Raymond, C. Schilero, K. Smolenski, F. Williams, C. Morrison, J. A. Borgia, M. J. Liptay, M. Pool, C. W. Seder, K. Junker, L. Omberg, M. Dinkin, G. Manikhas, D. Alvaro, M. C. Bragazzi, V. Cardinale, G. Carpino, E. Gaudio, D. Chesla, S. Cottingham, M. Dubina, F. Moiseenko, R. Dhanasekaran, K.-F. Becker, K.-P. Janssen, J. Slotta-Huspenina, M. H. Abdel-Rahman, D. Aziz, S. Bell, C. M. Cebulla, A. Davis, R. Duell, J. B. Elder, J. Hilty, B. Kumar, J. Lang, N. L. Lehman, R. Mandt, P. Nguyen, R. Pilarski, K. Rai, L. Schoenfield, K. Senecal, P. Wakely, P. Hansen, R. Lechan, J. Powers, A. Tischler, W. E. Grizzle, K. C. Sexton, A. Kastl, J. Henderson, S. Porten, J. Waldmann, M. Fassnacht, S. L. Asa, D. Schadendorf, M. Couce, M. Graefen, H. Huland, G. Sauter, T. Schlomm, R. Simon, P. Tennstedt, O. Olabode, M. Nelson, O. Bathe, P. R. Carroll, J. M. Chan, P. Disaia, P. Glenn, R. K. Kelley, C. N. Landen, J. Phillips, M. Prados, J. Simko, K. Smith-McCune, S. VandenBerg, K. Roggin, A. Fehrenbach, A. Kendler, S. Sifiri, R. Steele, A. Jimeno, F. Carey, I. Forgie, M. Mannelli, M. Carney, B. Hernandez, B. Campos, C. Herold-Mende, C. Jungk, A. Unterberg, A. von Deimling, A. Bossler, J. Galbraith, L. Jacobus, M. Knudson, T. Knutson, D. Ma, M. Milhem, R. Sigmund, A. K. Godwin, R. Madan, H. G. Rosenthal, C. Adebamowo, S. N. Adebamowo, A. Boussioutas, D. Beer, T. Giordano, A.-M. Mes-Masson, F. Saad, T. Bocklage, L. Landrum, R. Mannel, K. Moore, K. Moxley, R. Postier, J. Walker, R. Zuna, M. Feldman, F. Valdivieso, R. Dhir, J. Luketich, E. M. M. Pinero, M. Quintero-Aguilo, C. G. Carlotti, Jr., J. S. Dos Santos, R. Kemp, A. Sankarankuty, D. Tirapelli, J. Catto, K. Agnew, E. Swisher, J. Creaney, B. Robinson, C. S. Shelley, E. M. Godwin, S. Kendall, C. Shipman, C. Bradford, T. Carey, A. Haddad, J. Moyer, L. Peterson, M. Prince, L. Rozek, G. Wolf, R. Bowman, K. M. Fong, I. Yang, R. Korst, W. K. Rathmell, J. L. Fantacone-Campbell, J. A. Hooke, A. J. Kovatich, C. D. Shriver, J. DiPersio, B. Drake, R. Govindan, S. Heath, T. Ley, B. Van Tine, P. Westervelt, M. A. Rubin, J. I. Lee, N. D. Aredes, A. Mariamidze, J. M. Stuart, C. C. Benz and P. W. Laird, *Cell*, 2018, **173**, 291-304.
17. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. Fitzhugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczy, R. Levine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chisoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

- Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Correll, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, M. L. Hong, J. Dubois, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. De La Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordtsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. R. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. A. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, A. Patrinos and M. J. Morgan, *Nature*, 2001, **409**, 860-921.
18. D. Hanahan and R. A. Weinberg, *Cell*, 2000, **100**, 57-70.
 19. C. Lengauer, K. W. Kinzler and B. Vogelstein, *Nature*, 1998, **396**, 643-649.
 20. P. J. Hastings, J. R. Lupski, S. M. Rosenberg and G. Ira, *Nature Reviews Genetics*, 2009, **10**, 551-564.
 21. P. Hasty and C. Montagna, *Molecular & Cellular Oncology*, 2014, **1**, 1-12.
 22. M. W. N. Deininger, J. M. Goldman and J. V. Melo, *Blood*, 2000, **96**, 3343-3356.
 23. J. S. Butel, *Carcinogenesis*, 2000, **21**, 405-426.
 24. E. D. Pleasance, P. J. Stephens, S. O'Meara, D. J. McBride, A. Meynert, D. Jones, M.-L. Lin, D. Beare, K. W. Lau, C. Greenman, I. Varela, S. Nik-Zainal, H. R. Davies, G. R. Ordoñez, L. J. Mudie, C. Latimer, S. Edkins, L. Stebbings, L. Chen, M. Jia, C. Leroy, J. Marshall, A. Menzies, A. Butler, J. W. Teague, J. Mangion, Y. A. Sun, S. F. McLaughlin, H. E. Peckham, E. F. Tsung, G. L. Costa, C. C. Lee, J. D. Minna, A. Gazdar, E. Birney, M. D. Rhodes, K. J. McKernan, M. R. Stratton, P. A. Futreal and P. J. Campbell, *Nature*, 2009, **463**, 184-191.
 25. E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman, I. Varela, M.-L. Lin, G. R. Ordoñez, G. R. Bignell, K. Ye, J. Alipaz, M. J. Bauer, D. Beare, A. Butler, R. J. Carter, L. Chen, A. J. Cox, S. Edkins, P. I. Kokko-Gonzales, N. A. Gormley, R. J. Grocock, C. D. Haudenschild, M. M. Hims, T. James, M. Jia, Z. Kingsbury, C. Leroy, J. Marshall, A. Menzies, L. J. Mudie, Z. Ning, T. Royce, O. B. Schulz-Trieglaff, A. Spiridou, L. A. Stebbings, L. Szajkowski, J. Teague, D. Williamson, L. Chin, M. T. Ross, P. J. Campbell, D. R. Bentley, P. A. Futreal and M. R. Stratton, *Nature*, 2009, **463**, 191-197.
 26. I. D'Agostino-Jack and A. T. Shaw, *Nature Reviews Clinical Oncology*, 2017, **15**, 1-15.
 27. M. J. Garnett, E. J. Edelman, S. J. Heidorn, C. D. Greenman, A. Dastur, K. W. Lau, P. Greninger, I. R. Thompson, X. Luo, J. Soares, Q. Liu, F. Iorio, D. Surdez, L. Chen, R. J. Milano, G. R. Bignell, A. T. Tam, H. Davies, J. A. Stevenson, S. Barthorpe, S. R. Lutz, F. Kogera, K. Lawrence, A. McLaren-Douglas, X. Mitropoulos, T. Mironenko, H. Thi, L. Richardson, W. Zhou, F. Jewitt, T. Zhang, P. O'Brien, J. L. Boisvert, S. Price, W. Hur, W. Yang, X. Deng, A. Butler, H. G. Choi, J. W. Chang, J. Baselga, I. Stamenkovic, J. A. Engelman, S. V. Sharma, O. Delattre, J. Saez-Rodriguez, N. S. Gray, J. Settleman, P. A. Futreal, D. A. Haber, M. R. Stratton, S. Ramaswamy, U. McDermott and C. H. Benes, *Nature*, 2012, **483**, 570-577.
 28. S. Kobayashi, T. J. Boggon, T. Dayaram, P. A. Jänne, O. Kocher, M. Meyerson, B. E. Johnson, M. J. Eck, D. G. Tenen and B. Halmos, *New England Journal of Medicine*, 2005, **352**, 786-792.
 29. K. S. Thress, C. P. Paweletz, E. Felip, B. C. Cho, D. Stetson, B. Dougherty, Z. Lai, A. Markovets, A. Vivancos, Y. Kuang, D. Ercan, S. E. Matthews, M. Cantarini, J. C. Barrett, P. A. Jänne and G. R. Oxnard, *Nature Medicine*, 2015, **21**, 1-5.
 30. M. R. Stratton, P. J. Campbell and P. A. Futreal, *Nature*, 2009, **458**, 719-724.
 31. B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz and K. W. Kinzler, *Science*, 2013, **339**, 1546-1558.
 32. M. H. Bailey, C. Tokheim, E. Porta-Pardo, S. Sengupta, D. Bertrand, A. Weerasinghe, A. Colaprico, M. C. Wendl, J. Kim, B. Reardon, P. K.-S. Ng, K. J. Jeong, S. Cao, Z. Wang, J. Gao, Q. Gao, F. Wang, E. M. Liu, L. Mularoni, C. Rubio-Perez, N. Nagarajan, I. Cortés-Ciriano, D. C. Zhou, W.-W. Liang, J. M. Hess, V. D. Yellapantula, D. Tamborero, A. Gonzalez-Perez, C. Suphailai, J. Y. Ko, E. Khurana, P. J. Park, E. M. Van

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

Allen, H. Liang, S. J. Caesar-Johnson, J. A. Demchok, I. Felau, M. Kasapi, M. L. Ferguson, C. M. Hutter, H. J. Sofia, R. Tarnuzzer, Z. Wang, L. Yang, J. C. Zenklusen, J. Zhang, S. Chudamani, J. Liu, L. Lolla, R. Naresh, T. Pihl, Q. Sun, Y. Wan, Y. Wu, J. Cho, T. DeFreitas, S. Frazer, N. Gehlenborg, G. Getz, D. I. Heiman, J. Kim, M. S. Lawrence, P. Lin, S. Meier, M. S. Noble, G. Saksena, D. Voet, H. Zhang, B. Bernard, N. Chambwe, V. Dhankani, T. Knijnenburg, R. Kramer, K. Leinonen, Y. Liu, M. Miller, S. Reynolds, I. Shmulevich, V. Thorsson, W. Zhang, R. Akbani, B. M. Broom, A. M. Hegde, Z. Ju, R. S. Kanchi, A. Korkut, J. Li, H. Liang, S. Ling, W. Liu, Y. Lu, G. B. Mills, K.-S. Ng, A. Rao, M. Ryan, J. Wang, J. N. Weinstein, J. Zhang, A. Abeshouse, J. Armenia, D. Chakravarty, W. K. Chatila, I. de Bruijn, J. Gao, B. E. Gross, Z. J. Heins, R. Kundra, K. La, M. Ladanyi, A. Luna, M. G. Nissan, A. Ochoa, S. M. Phillips, E. Reznik, F. Sanchez-Vega, C. Sander, N. Schultz, R. Sheridan, S. O. Sumer, Y. Sun, B. S. Taylor, J. Wang, H. Zhang, P. Anur, M. Peto, P. Spellman, C. Benz, J. M. Stuart, C. K. Wong, C. Yau, D. N. Hayes, J. S. Parker, M. D. Wilkerson, A. Ally, M. Balasundaram, R. Bowlby, D. Brooks, R. Carlsen, E. Chuah, N. Dhalla, R. Holt, S. J. M. Jones, K. Kasaian, D. Lee, Y. Ma, M. A. Marra, M. Mayo, R. A. Moore, A. J. Mungall, K. Mungall, A. G. Robertson, S. Sadeghi, J. E. Schein, P. Sipahimalani, A. Tam, N. Thiessen, K. Tse, T. Wong, A. C. Berger, R. Beroukhi, A. D. Cherniack, C. Cibulskis, S. B. Gabriel, G. F. Gao, G. Ha, M. Meyerson, S. E. Schumacher, J. Shih, M. H. Kucherlapati, R. S. Kucherlapati, S. Baylin, L. Cope, L. Danilova, M. S. Bootwalla, P. H. Lai, D. T. Maglinte, D. J. Van Den Berg, D. J. Weisenberger, J. T. Auman, S. Balu, T. Bodenheimer, C. Fan, K. A. Hoadley, A. P. Hoyle, S. R. Jefferys, C. D. Jones, S. Meng, P. A. Mieczkowski, L. E. Mose, A. H. Perou, C. M. Perou, J. Roach, Y. Shi, J. V. Simons, T. Skelly, M. G. Soloway, D. Tan, U. Veluvolu, H. Fan, T. Hinoue, P. W. Laird, H. Shen, W. Zhou, M. Bellair, K. Chang, K. Covington, C. J. Creighton, H. Dinh, H. Doddapaneni, L. A. Donehower, J. Drummond, R. A. Gibbs, R. Glenn, W. Hale, Y. Han, J. Hu, V. Korchina, S. Lee, L. Lewis, W. Li, X. Liu, M. Morgan, D. Morton, D. Muzny, J. Santibanez, M. Sheth, E. Shinbrot, L. Wang, M. Wang, D. A. Wheeler, L. Xi, F. Zhao, J. Hess, E. L. Appelbaum, M. Bailey, M. G. Cordes, L. Ding, C. C. Fronick, L. A. Fulton, R. S. Fulton, C. Kandoth, E. R. Mardis, M. D. McLellan, C. A. Miller, H. K. Schmidt, R. K. Wilson, D. Crain, E. Curley, J. Gardner, K. Lau, D. Mallery, S. Morris, J. Paulauskis, R. Penny, C. Shelton, T. Shelton, M. Sherman, E. Thompson, P. Yena, J. Bowen, J. M. Gastier-Foster, M. Gerken, K. M. Leraas, T. M. Lichtenberg, N. C. Ramirez, L. Wise, E. Zmuda, N. Corcoran, T. Costello, C. Hovens, A. L. Carvalho, A. C. de Carvalho, J. H. Fregani, A. Longatto-Filho, R. M. Reis, C. Scapulatempo-Neto, H. C. S. Silveira, D. O. Vidal, A. Burnette, J. Eschbacher, B. Hermes, A. Noss, R. Singh, M. L. Anderson, P. D. Castro, M. Ittmann, D. Huntsman, B. Kohl, X. Le, R. Thorp, C. Andry, E. R. Duffy, V. Lyadov, O. Paklina, G. Setdikova, A. Shabunin, M. Tavobilov, C. McPherson, R. Warnick, R. Berkowitz, D. Cramer, C. Feltmate, N. Horowitz, A. Kibel, M. Muto, C. P. Raut, A. Malykh, J. S. Barnholtz-Sloan, W. Barrett, K. Devine, J. Fulop, Q. T. Ostrom, K. Shimmel, Y. Wolinsky, A. E. Sloan, A. De Rose, F. Giuliante, M. Goodman, B. Y. Karlan, C. H. Hagedorn, J. Eckman, J. Harr, J. Myers, K. Tucker, L. A. Zach, B. Deyarmin, H. Hu, L. Kvecher, C. Larson, R. J. Mural, S. Somiari, A. Vicha, T. Zelinka, J. Bennett, M. Iacocca, B. Rabeno, P. Swanson, M. Latour, L. Lacombe, B. Têtu, A. Bergeron, M. McGraw, S. M. Staugaitis, J. Chabot, H. Hibshoosh, A. Sepulveda, T. Su, T. Wang, O. Potapova, O. Voronina, L. Desjardins, O. Mariani, S. Roman-Roman, X. Sastre, M.-H. Stern, F. Cheng, S. Signoretti, A. Berchuck, D. Bigner, E. Lipp, J. Marks, S. McCall, R. McLendon, A. Secord, A. Sharp, M. Behera, D. J. Brat, A. Chen, K. Delman, S. Force, F. Khuri, K. Magliocca, S. Maittel, J. J. Olson, T. Owonikoko, A. Pickens, S. Ramalingam, D. M. Shin, G. Sica, E. G. Van Meir, H. Zhang, W. Eijckenboom, A. Gillis, E. Korpershoek, L. Looijenga, W. Oosterhuis, H. Stoop, K. E. van Kessel, E. C. Zwarthoff, C. Calatozzolo, L. Cuppini, S. Cuzzubbo, F. DiMeco, G. Finocchiaro, L. Mattei, A. Perin, B. Pollo, C. Chen, J. Houck, P. Lohavanichbutr, A. Hartmann, C. Stoehr, R. Stoehr, H. Taubert, S. Wach, B. Wullich, W. Kyler, D. Murawa, M. Wiznerowicz, K. Chung, W. J. Edenfield, J. Martin, E. Baudin, G. Bublely, R. Bueno, A. De Rienzo, W. G. Richards, S. Kalkanis, T. Mikkelsen, H. Noushmehr, L. Scarpace, N. Girard, M. Aymerich, E. Campo, E. Giné, A. L. Guillermo, N. Van Bang, P. T. Hanh, B. D. Phu, Y. Tang, H. Colman, K. Evason, P. R. Dottino, J. A. Martignetti, H. Gabra, H. Juhl, T. Akeredolu, S. Stepa, D. Hoon, K. Ahn, K. J. Kang, F. Beuschlein, A. Breggia, M. Birrer, D. Bell, M. Borad, A. H. Bryce, E. Castle, V. Chandan, J. Cheville, J. A. Copland, M. Farnell, T. Flotte, N. Giama, T. Ho, M. Kendrick, J.-P. Kocher, K. Kopp, C. Moser, D. Nagorney, D. O'Brien, B. P. O'Neill, T. Patel, G. Petersen, F. Que, M. Rivera, L. Roberts, R. Smallridge, T. Smyrk, M. Stanton, R. H. Thompson, M. Torbenson, J. D. Yang, L. Zhang, F. Brimo, J. A. Ajani, A. M. A. Gonzalez, C. Behrens, J. Bondaruk, R. Broaddus, B. Czerniak, B. Esmaeli, J. Fujimoto, J. Gershenwald, C. Guo, A. J. Lazar, C. Logothetis, F. Meric-Bernstam, C. Moran, L. Ramondetta, D. Rice, A. Sood, P. Tamboli, T. Thompson, P. Troncso, A. Tsao, I. Wistuba, C. Carter, L. Haydu, P. Hersey, V. Jakrot, H. Kakavand, R. Kefford, K. Lee, G. Long, G. Mann, M. Quinn, R. Saw, R. Scolyer, K. Shannon, A. Spillane, J. Stretch, M. Synott, J. Thompson, J. Wilmott, H. Al-Ahmadie, T. A. Chan, R. Ghossein, A. Gopalan, D. A. Levine, V. Reuter, S. Singer, B. Singh, N. V. Tien, T. Broudy, C. Mirsaiidi, P. Nair, P. Drwiega, J. Miller, J. Smith, H. Zaren, J.-W. Park, N. P.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

- Hung, E. Kebebew, W. M. Linehan, A. R. Metwalli, K. Pacak, P. A. Pinto, M. Schiffman, L. S. Schmidt, C. D. Vocke, N. Wentzensen, R. Worrell, H. Yang, M. Moncrieff, C. Goparaju, J. Melamed, H. Pass, N. Botnariuc, I. Caraman, M. Cernat, I. Chemencedji, A. Clipca, S. Doruc, G. Gorincioi, S. Mura, M. Pirtac, I. Stancul, D. Tcaciuc, M. Albert, I. Alexopoulou, A. Arnaout, J. Bartlett, J. Engel, S. Gilbert, J. Parfitt, H. Sekhon, G. Thomas, D. M. Rassl, R. C. Rintoul, C. Bifulco, R. Tamakawa, W. Urba, N. Hayward, H. Timmers, A. Antenucci, F. Facciolo, G. Grazi, M. Marino, R. Merola, R. de Krijger, A.-P. Gimenez-Roqueplo, A. Piché, S. Chevalier, G. McKercher, K. Birsoy, G. Barnett, C. Brewer, C. Farver, T. Naska, N. A. Pennell, D. Raymond, C. Schilero, K. Smolenski, F. Williams, C. Morrison, J. A. Borgia, M. J. Liptay, M. Pool, C. W. Seder, K. Junker, L. Omberg, M. Dinkin, G. Manikhas, D. Alvaro, M. C. Bragazzi, V. Cardinale, G. Carpino, E. Gaudio, D. Chesla, S. Cottingham, M. Dubina, F. Moiseenko, R. Dhanasekaran, K.-F. Becker, K.-P. Janssen, J. Slotta-Huspenina, M. H. Abdel-Rahman, D. Aziz, S. Bell, C. M. Cebulla, A. Davis, R. Duell, J. B. Elder, J. Hilty, B. Kumar, J. Lang, N. L. Lehman, R. Mandt, P. Nguyen, R. Pilarski, K. Rai, L. Schoenfield, K. Senecal, P. Wakely, P. Hansen, R. Lechan, J. Powers, A. Tischler, W. E. Grizzle, K. C. Sexton, A. Kastl, J. Henderson, S. Porten, J. Waldmann, M. Fassnacht, S. L. Asa, D. Schadendorf, M. Couce, M. Graefen, H. Huland, G. Sauter, T. Schlomm, R. Simon, P. Tennstedt, O. Olabode, M. Nelson, O. Bathe, P. R. Carroll, J. M. Chan, P. Disaia, P. Glenn, R. K. Kelley, C. N. Landen, J. Phillips, M. Prados, J. Simko, K. Smith-McCune, S. VandenBerg, K. Roggin, A. Fehrenbach, A. Kendler, S. Sifri, R. Steele, A. Jimeno, F. Carey, I. Forgie, M. Mannelli, M. Carney, B. Hernandez, B. Campos, C. Herold-Mende, C. Jungk, A. Unterberg, A. von Deimling, A. Bossler, J. Galbraith, L. Jacobus, M. Knudson, T. Knutson, D. Ma, M. Milhem, R. Sigmund, A. K. Godwin, R. Madan, H. G. Rosenthal, C. Adebamowo, S. N. Adebamowo, A. Boussioutas, D. Beer, T. Giordano, A.-M. Mes-Masson, F. Saad, T. Bocklage, L. Landrum, R. Mannel, K. Moore, K. Moxley, R. Postier, J. Walker, R. Zuna, M. Feldman, F. Valdivieso, R. Dhir, J. Luketich, E. M. M. Pinero, M. Quintero-Aguilo, C. G. Carlotti, J. S. Dos Santos, R. Kemp, A. Sankarankuty, D. Tirapelli, J. Catto, K. Agnew, E. Swisher, J. Creaney, B. Robinson, C. S. Shelley, E. M. Godwin, S. Kendall, C. Shipman, C. Bradford, T. Carey, A. Haddad, J. Moyer, L. Peterson, M. Prince, L. Rozek, G. Wolf, R. Bowman, K. M. Fong, I. Yang, R. Korst, W. K. Rathmell, J. L. Fantacone-Campbell, J. A. Hooke, A. J. Kovatich, C. D. Shriver, J. DiPersio, B. Drake, R. Govindan, S. Heath, T. Ley, B. Van Tine, P. Westervelt, M. A. Rubin, J. I. Lee, N. D. Aredes, A. Mariamidze, M. S. Lawrence, A. Godzik, N. Lopez-Bigas, J. Stuart, D. Wheeler, G. Getz, K. Chen, A. J. Lazar, G. B. Mills, R. Karchin and L. Ding, *Cell*, 2018, **173**, 371-385.
33. C. M. Croce, *New England Journal of Medicine*, 2008, **358**, 502-511.
 34. J. Downward, *Nature Reviews Cancer*, 2003, **3**, 11-22.
 35. C.-H. Pui, M. V. Relling and J. R. Downing, *New England Journal of Medicine*, 2004, **350**, 1535-1548.
 36. J. M. Bishop, *Cell*, 1991, **64**, 235-248.
 37. E. Y. H. P. Lee and W. J. Muller, *Cold Spring Harbor Perspectives in Biology*, 2010, **2**, 1-18.
 38. I. A. Prior, P. D. Lewis and C. Mattos, *Cancer Research*, 2012, **72**, 2457-2467.
 39. H. Chen, H. Liu and G. Qing, *Signal Transduction and Targeted Therapy*, 2018, **3**, 1-7.
 40. D. W. Felsher, *Nature Reviews Cancer*, 2003, **3**, 375-380.
 41. C. J. Sherr, *Cell*, 2004, **116**, 235-246.
 42. A. G. Knudson, *Proceedings of the National Academy of Sciences*, 1971, **68**, 820-823.
 43. F. Michor, Y. Iwasa and M. A. Nowak, *Nature Reviews Cancer*, 2004, **4**, 197-205.
 44. A. J. Levine, *Cell*, 1997, **88**, 323-331.
 45. S. Surget, M. P. Khoury and J.-C. Bourdon, *OncoTargets and Therapy*, 2014, **7**, 57-68.
 46. S. P. Jackson and J. Bartek, *Nature*, 2009, **461**, 1071-1078.
 47. M. B. Kastan and J. Bartek, *Nature*, 2004, **432**, 316-323.
 48. J. H. J. Hoeijmakers, *Nature*, 2001, **411**, 366-374.
 49. J. D. Fackenthal and O. I. Olopade, *Nature Reviews Cancer*, 2007, **7**, 937-948.
 50. P. A. Jones and S. B. Baylin, *Nature Reviews Genetics*, 2002, **3**, 415-428.
 51. P. A. Jones and S. B. Baylin, *Cell*, 2007, **128**, 683-692.
 52. A. Esquela-Kerscher and F. J. Slack, *Nature Reviews Cancer*, 2006, **6**, 259-269.
 53. A. A. Svoronos, D. M. Engelman and F. J. Slack, *Cancer Research*, 2016, **76**, 3666-3670.
 54. L. Chin, J. N. Andersen and P. A. Futreal, *Nature Medicine*, 2011, **17**, 297-303.
 55. L. M. Coussens and Z. Werb, *Nature*, 2002, **420**, 860-867.
 56. S. M. Crusz and F. R. Balkwill, *Nature Reviews Clinical Oncology*, 2015, **12**, 1-13.
 57. S. I. Grivnikov, F. R. Greten and M. Karin, *Cell*, 2010, **140**, 883-899.
 58. P. Carmeliet and R. K. Jain, *Nature*, 2000, **407**, 249-257.
 59. P. Carmeliet and R. K. Jain, *Nature*, 2011, **473**, 298-307.
 60. D. Wirtz, K. Konstantopoulos and P. C. Searson, *Nature Reviews Cancer*, 2011, **11**, 512-522.

Chapter 1 – The Burden of Cancer, Cancer Biology and Targeted Treatment

61. D. F. Quail and J. A. Joyce, *Nature Medicine*, 2013, **19**, 1423-1437.
62. M. G. Vander Heiden, L. C. Cantley and C. B. Thompson, *Science*, 2009, **324**, 1029-1033.
63. J. B. Swann and M. J. Smyth, *The Journal of Clinical Investigation*, 2007, **117**, 1137-1146.
64. C. G. Drake, E. Jaffee and D. M. Pardoll, *Advances in Immunology*, 2006, **90**, 51-81.
65. D. S. Vinay, E. P. Ryan, G. Pawelec, W. H. Talib, J. Stagg, E. Elkord, T. Lichtor, W. K. Decker, R. L. Whelan, H. M. C. S. Kumara, E. Signori, K. Honoki, A. G. Georgakilas, A. Amin, W. G. Helferich, C. S. Boosani, G. Guha, M. R. Ciriolo, S. Chen, S. I. Mohammed, A. S. Azmi, W. N. Keith, A. Bilsland, D. Bhakta, D. Halicka, H. Fujii, K. Aquilano, S. S. Ashraf, S. Nowsheen, X. Yang, B. K. Choi and B. S. Kwon, *Seminars in Cancer Biology*, 2015, **35**, S185-S198.
66. K. Strebhardt and A. Ullrich, *Nature Reviews Cancer*, 2008, **8**, 473-480.
67. V. T. DeVita and E. Chu, *Cancer Research*, 2008, **68**, 8643-8653.
68. B. A. Chabner and T. G. Roberts Jr, *Nature Reviews Cancer*, 2005, **5**, 65-72.
69. R. Capdeville, E. Buchdunger, J. Zimmermann and A. Matter, *Nature Reviews Drug Discovery*, 2002, **1**, 493-502.
70. B. Gyawali and R. Sullivan, *The New Bioethics*, 2017, **23**, 95-104.
71. C. R. Chong and P. A. Jänne, *Nature Medicine*, 2013, **19**, 1389-1400.
72. D. Buffery, *American Health and Drug Benefits*, 2016, **9**, 233-238.
73. J. Sun, Q. Wei, Y. Zhou, J. Wang, Q. Liu and H. Xu, *BMC Systems Biology*, 2017, **11**, 27-43.
74. S. Mayor, *The Lancet Oncology*, 2015, **16**, 757-758.
75. G. Bergers and D. Hanahan, *Nature Reviews Cancer*, 2008, **8**, 592-603.
76. J. C. Reed, *Cancer Cell*, 2003, **3**, 17-22.
77. D. M. Pardoll, *Nature Reviews Cancer*, 2012, **12**, 252-264.
78. Mark A. Dawson and T. Kouzarides, *Cell*, 2012, **150**, 12-27.

Chapter 2

Epidermal Growth Factor Receptor and its Targeted Inhibition

Abstract

This chapter will introduce EGFR as the therapeutic target of relevance for this PhD thesis - from its discovery in the 1950s to its targeting in clinical treatment of NSCLC. A description of the structural makeup and receptor family, as well as its significance in the regulation of cellular process, will give insight into the mechanism of its activation and its complex network of downstream signalling cascades. Discussion of the factors which promote oncogenic signalling of these pathways, such as overexpression and hyperactivation, will serve to justify its selection, and inhibition thereof, in the treatment of various cancers. To gain an understanding of the targeted inhibition of EGFR, protein kinases and their central role in reversible phosphorylation will be considered, as the catalytic active site of this process forms the main binding site for the majority of inhibitors. The various types, binding modes and means of inhibition of these drugs candidates will be examined. Lastly, the history of EGFR as a targeted therapeutic will be discussed in detail, from the 1st generation of inhibitors to the present 4th generation. Additionally, critical successes and failures over the years, as well as future approaches needed to overcome the inevitable rise of drug resistance mutations within the receptor, will be analysed within this chapter.

2.1 Epidermal Growth Factor Receptor

EGFR is a transmembrane receptor which belongs to the ErbB/HER family of protein-tyrosine kinases. Since its discovery by Stanley Cohen, it has become one of the most studied receptors in biology, owing to its expansive role in signal transduction pathways.¹ Through these signalling cascades, a host of essential aspects of cellular life are regulated, including division, growth and apoptosis.² EGFR was the first receptor to provide evidence of a relationship between receptor overexpression and cancer, one of the various mechanisms that lead to oncogenic signalling and hyperactivation.³ In this section, we will discuss the uncovering of this receptor, its structural makeup, the machinery and mechanisms of the ErbB family signalling network, and lastly, the factors that lead to its oncogenic activation.

2.1.1 Discovery

The 1986 Nobel Prize in Physiology and Medicine was awarded to the Italian developmental biologist Rita Levi-Montalcini and the American biochemist Stanley Cohen for their discovery of nerve growth factor (NGF) and epidermal growth factor (EGF), respectively.^{4,5} These discoveries were fundamentally important in our understanding of the mechanisms which regulate cell and organ growth and were the first of many growth regulating substances to be discovered and characterised.

Scientists in the 1950s did not understand how the development of the nervous system resulted in the complete innervation of the human body or how it was regulated. It was Levi-Montalcino, shown in **Figure 2.1a**, who first observed that when certain mouse tumours were implanted into chick embryos, they induced potent growth of the embryonic nervous system - specifically the sensory and sympathetic neurons.⁵ It was concluded that the tumour released a nerve growth-promoting factor which had a selective action on certain neurons.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

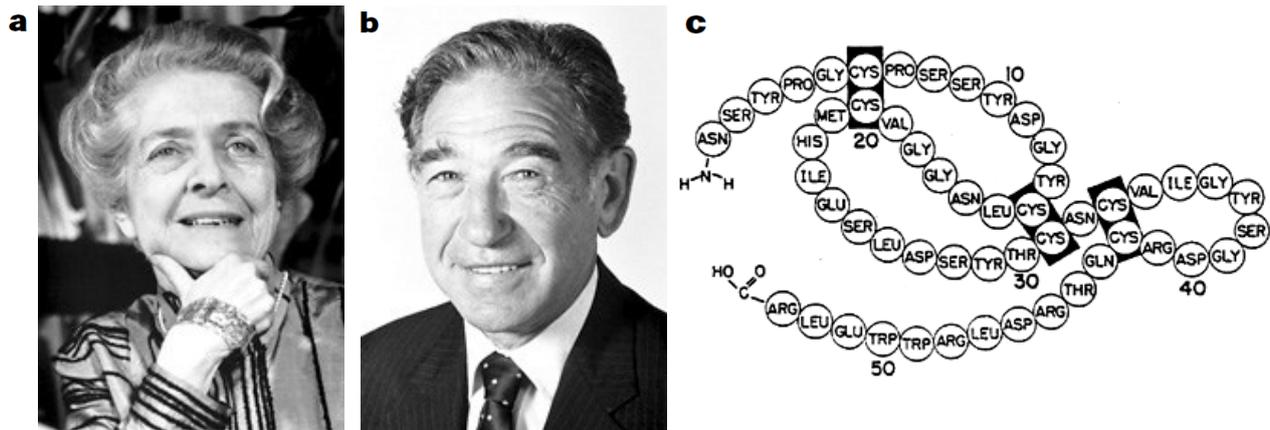


Figure 2.1: The recipients of the Nobel Prize in Physiology and Medicine in 1986 a) Rita Levi-Montalcini, b) Stanley Cohen and c) Cohen's structure of EGF. Figure reproduced from c) Cohen.⁵

The development of a biological assay to detect NGF by Levi-Montalcino paved the way for Cohen, shown in **Figure 2.1b**, to identify a rich source of the growth factor – the salivary gland in a male mouse. Experiments with snake venom and the extract from the salivary gland led to the purification and isolation of NGF and its antibodies. From this work, the chemical structure of NGF was elucidated – a milestone in developmental neurobiology.⁵

During his research on NGF, Cohen observed unexpected side-effects not related to the activities of nerve growth factor when injecting the salivary gland extract into new-born mice. The new-born mice underwent an acceleration of development, displaying precocious eyelid opening (6-7 days, in comparison to 12-14 days for the control group) and tooth eruption (5-6 days, in comparison to 8-10 days for the control group). Cohen postulated the existence of another growth factor within the extract, and termed this substance epidermal growth factor, as it stimulated the growth of epithelial cells in the cornea and skin. As with NGF, Cohen raised antibodies against, isolated and purified EGF, leading to the determination of the amino acid sequence and structure, as can be seen in **Figure 2.1c**.⁵

For the first time, scientists had access to a factor which stimulated epithelial cell growth, and this allowed Cohen to investigate the growth process. He found that EGF initiated a cascade of events which stimulated protein and DNA synthesis, cell division and growth of a large variety of cells. A prerequisite for the action of EGF was the existence of binding sites, termed receptors, which were found on the surface of target cells. The epidermal growth factor receptor (EGFR) was thus described to catch EGF and form a complex, activating the signalling cascade.

The phosphorylation of the amino acid tyrosine on this receptor was found to be crucial to its biological action and was later found to be a general pathway through which many growth factors mediate their effects. Knowledge of the regulation of normal cell growth led to an understanding of the mechanisms of the aberrant and uncontrolled cellular growth of tumours. The structure and detailed mechanisms involved in the signalling pathway, as well as its oncogenic activation, will be discussed in the following sections.

2.1.2 Receptor Structure and Ligand Binding

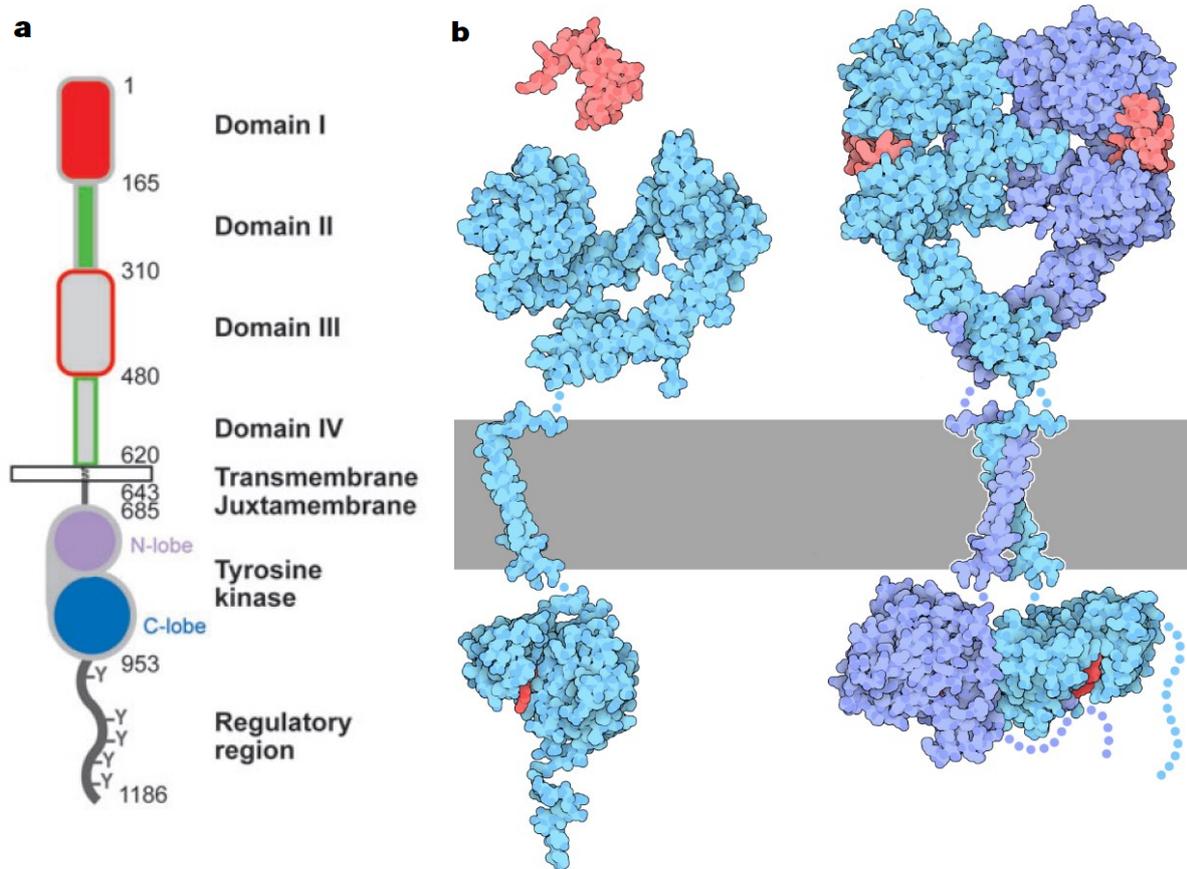


Figure 2.2: a) The domains of EGFR with corresponding amino acid residues and b) The inactive (left) and active dimer (right) of EGFR, with ligands shown in red and cell membrane in grey. Figures reproduced from a) Ferguson et al and b) Goodsell.^{6,7}

The human EGFR family consists of four members that belong to the ErbB lineage of proteins.² These include and are designated as follows: i) EGFR/ERBB1/HER1, ii) ERBB2/HER2/NEU, iii) ERBB3/HER3 and iv) ERBB4/HER4. These homologous proteins all consist of a glycosylated extracellular domain, a single hydrophobic transmembrane segment and an intracellular domain that facilitates protein kinase activity, as shown in **Figure 2.2a**, with amino acid numbers denoting each domain boundary.⁶ The large extracellular region is divided into four domains. Domain I and III are leucine-rich segments that share approximately 37% sequence identity and participate in ligand binding. Domain II and IV are abundant in cysteine residues which allow for disulphide bond formation, with domain II partaking in homo- and heterodimer formation exclusively. The extracellular region is followed by a short transmembrane segment, leading to an intracellular portion of roughly 550 amino acid residues. This intracellular portion consists of a juxtamembrane segment, a protein tyrosine kinase domain and a long, flexible carboxyterminal tail.²

The ErbB family of proteins function as homo- and heterodimers. Dimerization is induced by binding of a specific ligand to the receptor site and is illustrated in **Figure 2.2b**.⁷ Prior to a ligand such as EGF (red) binding, the inactive form of the receptor (light blue) folds back on itself (**Figure 2.2b**, left). Successful ligand binding results in a conformational change and exposure of a dimerization arm in domain II of the extracellular portion. The two ligand-bound EGFR complexes (light and dark blue) then unite to form a back-to-back dimer, with the ligands positioned on opposing sides of the dimer (**Figure 2.2b**, right).

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

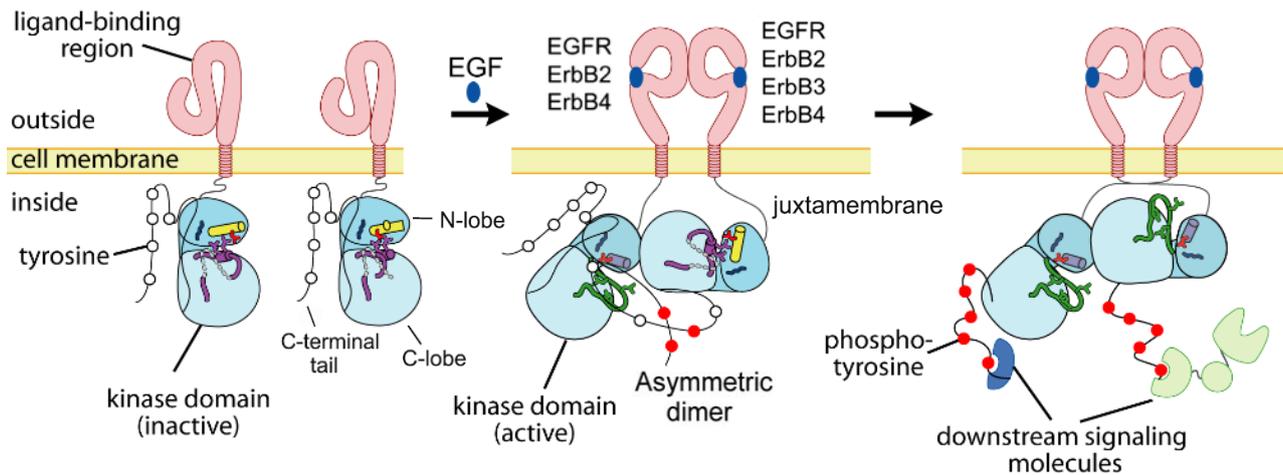


Figure 2.3: A general model for the activation of the intracellular kinase domain of ErbB family members. Figure redrawn from Zhang et al.⁸

Although EGFR was the first receptor tyrosine kinase to be discovered, and the structural and mechanistic basis for ligand-induced dimerization of its extracellular region has been well established, the mechanism of activation of its intracellular kinase domain was not clearly understood until recently. This is because its activation mechanism differs from other receptor kinases. The general mechanism of activation of most protein kinases is initiated by symmetric dimer formation and requires phosphorylation of one or more residues within the activation segment. This is accomplished in *trans*, i.e., the first member of the dimer mediates phosphorylation of the second and vice versa.² While this manner of phosphorylation of Tyr869 in the activation segment of EGFR does occur, this is not required for its activation.

Using mutational analysis and protein crystallography, it was found that the ligand-activated EGFR kinase domains form an asymmetric homodimer, which resembles that of the heterodimer formed by cyclin-dependant kinase 2 (CDK2) and its activating protein cyclin A.⁸ In this asymmetric homodimer, shown in the middle of **Figure 2.3**, one kinase domain plays the role of the activator/donor and the other the activated/acceptor. The amino-terminal lobe (N-lobe) of the activated/acceptor kinase interacts with the carboxyterminal lobe (C-lobe) of the donor kinase, leading to its allosteric stimulation and catalysing the phosphorylation of tyrosine residues on the C-terminal tail of the donor kinase (**Figure 2.3**, middle).⁸ The importance of the juxtamembrane segment of the receptor in this process has recently been recognized.⁹ The juxtamembrane potentiates kinase activity by latching the donor kinase domain to the acceptor, facilitating dimerization and phosphorylation.

All members of the ErbB family of enzymes make use of a similar mechanism of activation after ligand-induced homo- or heterodimerization. The phosphorylated tyrosine residues serve as docking sites for signalling molecules which enable downstream signalling. Following EGFR activation, a rich network of some of the most extensively studied signalling pathways are triggered, initiating a cascade of events that play an important role in cell division, death, motility and adhesion. These pathways, and their implication in cancer, will be analysed and discussed in the following sections.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

2.1.3 The ErbB Signalling Network

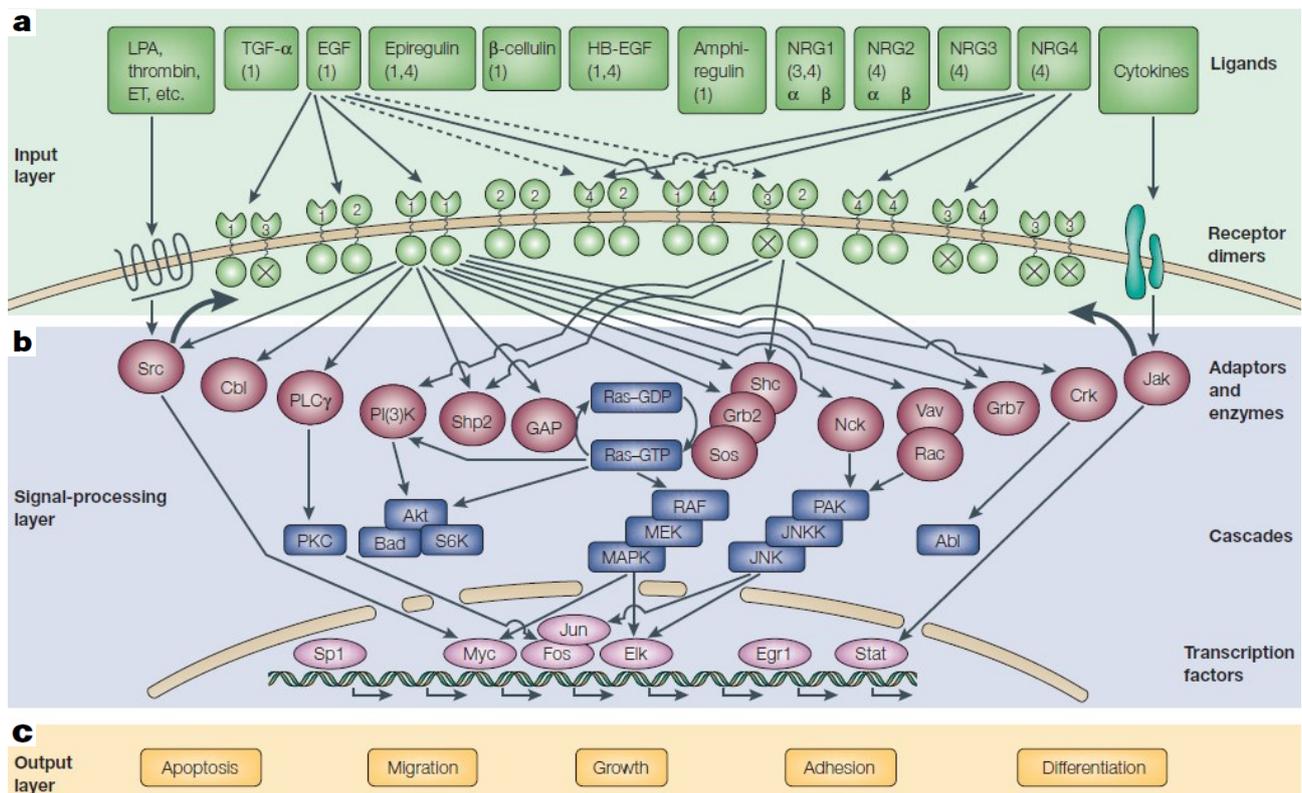


Figure 2.4: The ErbB family signalling network. Figure reproduced from Yarden and Sliwkowski.¹⁰

The components of the ErbB signalling pathways are evolutionarily ancient, with certain nematodes and fruit flies possessing primordial linear versions of this pathway. In higher eukaryotes, this simple pathway has evolved into a complex network, presumably due to the selective gains in terms of adaptation, mutation tolerance and diversification of signals conferred by an interconnected layered structure.¹¹ This richly interactive and multi-layered network, depicted in **Figure 2.4**, provides context-specific biological responses through combinatorial expression and activation of various components.¹⁰ The network can be divided into three simplified layers, namely, the input, signal processing and output layers.

2.1.3.1 Input Layer

The input layer (**Figure 2.4a**), comprises of the growth factors (ligands) and the ErbB family of receptors. There are 11 growth factors in the EGF-like family and the four ErbB family members are able to form 28 homo- and heterodimers, with 26 being active. This amounts to a total of 614 possible ligand-bound, homo- and heterodimer combinations of receptors, with the number being reduced to 611 after omitting non-functional homodimers.² EGFR is able to bind to 7 of the 11 growth factors, notably including EGF and transforming growth factor alpha (TGF- α). These growth factors are shown in **Figure 2.4a** as green squares, with the numbers in the boxes representing high affinity binding receptors. Receptor homo- and heterodimer combinations are shown with numbered partners, with only the specific pathways of receptors binding to EGF and neuregulin 4 (NG4) shown for simplicity. Ligand binding, followed by dimerization and kinase activation and phosphorylation, provide docking sites for various adaptor proteins or enzymes within the signalling-processing layer.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

2.1.3.2 Signal-processing layer

Several factors determine the specificity and potency of intracellular signalling pathways, with the main determinant being the vast array of phosphotyrosine binding proteins that interact with the phosphorylated C-tail molecules of each ErbB dimeric complex, shown in **Figure 2.4b** as purple circles.¹⁰ The phosphorylation site, and thus which of these enzymes are engaged, is determined by identity of the ligand, dimer composition and the structural makeup of the receptor in the input layer. For simplicity, only pathways to the signalling proteins from two dimeric complexes - the ErbB1 homodimer and the ErbB2-ErbB3 heterodimer - are shown in the signal-processing layer (**Figure 2.4b**).¹⁰ The interaction of these adaptor enzymes with the phosphotyrosine molecules triggers a variety of downstream signalling cascade pathways (**Figure 2.4b**, blue rectangles).¹⁰ Activation of these signalling modules translate in the nucleus of the cell into distinct transcriptional programs through transcription factors, (**Figure 2.4b**, pink ovals).¹⁰ These cascades are ultimately responsible for the correct output of core cellular processes such as cell growth, division and death. Only some of the pathways and transcription factors are represented in **Figure 2.4b**, with the three main downstream signalling modules being the i) RAS-RAF-MEK-MAPK, ii) PI3K-AKT and iii) PLC γ -PKC. The RAS-MAPK pathway is an invariable target of all ErbB ligands, highlighting its biological importance, and the PI3K-AKT pathway is downstream from the majority of ErbB dimers. The activation of these cascades, and regulation of transcription factors translating to a specific type of biological output, is critical to a cells normal functioning.

2.1.3.3 Output layer

The output of the ErbB signalling network, shown in **Figure 2.3c**, can range from cell motility and division, to differentiation, adhesion and programmed cellular death.² These specific outputs depend on the cellular context and, as discussed above, the particular ligand and ErbB dimer complex leading to signal processing. In terms of mitogenic and transforming outputs, homodimeric combinations of receptors are weaker than corresponding heterodimeric combinations, and heterodimeric combinations with ErbB2, such as EGFR-ErbB2, have been shown to be the most potent complexes.¹⁰ Following ligand binding and dimerization, the adaptor and effector proteins discussed in the signal-processing layer further stimulate the signalling cascades, leading to various cellular process outputs. The RAS-MAPK and PLC γ -PKC pathways participate in cell proliferation, whilst the PI3K-AKT pathway plays an important role in mediating cell survival. Other ErbB signalling modules not highlighted, participate in angiogenesis, cell adhesion, cell motility and organogenesis.² Given the functional importance and involvement of EGFR and its related family members in these diverse cellular processes, their aberrant activity and its inference in cancer is not surprising. Mutations harboured and gained in proto-oncogenes, which code for the relevant proteins involved in the signalling cascade, often leads to deregulation of these cellular processes and uncontrolled growth of malignant cells, leading to tumorigenesis. This oncogenic activation is implicated in a multitude of cancers and has warranted its relevance as a therapeutic target.

2.1.4 Oncogenic Activation of EGFR and Cancer

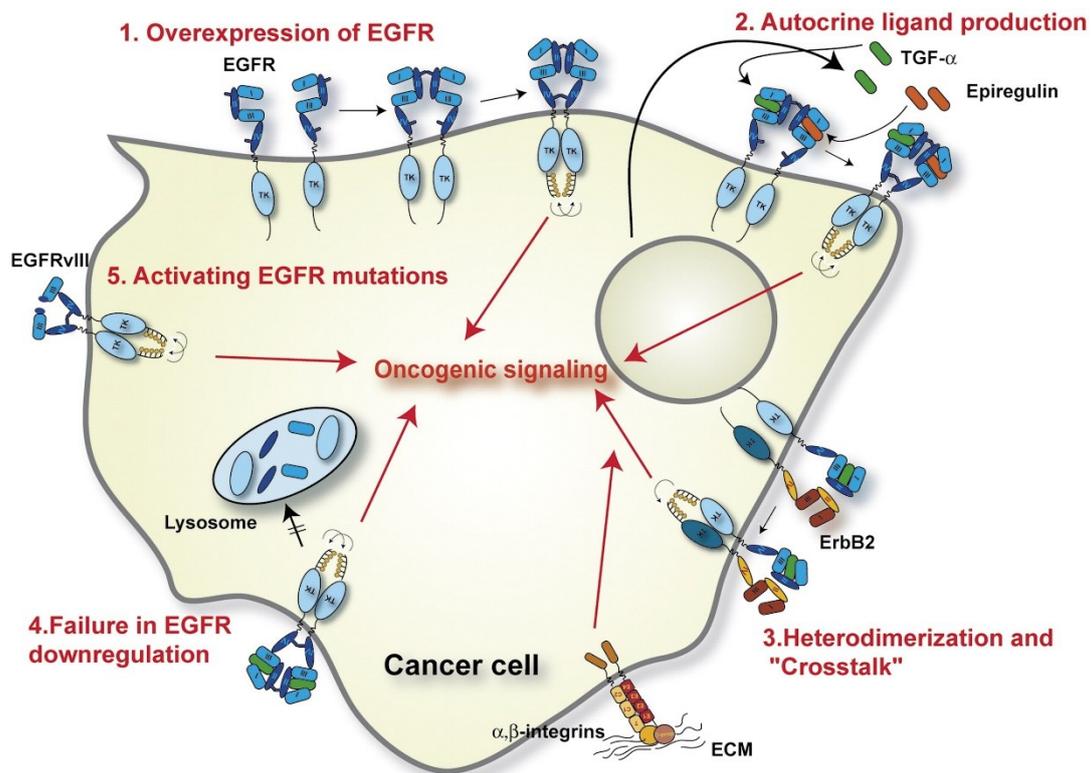


Figure 2.5: Mechanisms leading to oncogenic signalling in EGFR. Figure reproduced from Zandi et al.¹²

Various cancers are associated with the mutation or increased expression of members of the ErbB family of receptors and their respective ligands. These malignancies include lung, breast, stomach, colorectal, head and neck, pancreatic carcinomas and glioblastoma.² This is particularly prevalent in lung and breast cancer, with EGFR mutations occurring in up to 40% of lung cancer samples and being overexpressed in 60% of NSCLC tumours, and ErbB2/HER2 gene amplification or overexpression occurring in up to 30% of breast cancers.^{2, 13} Cancer cells make use of the potent cell proliferation signals generated by the ErbB network to carry over oncogenic mutations by clonal expansion. This is accomplished through several mechanisms of hyperactivation of ErbB pathways at different layers in the network, such as enhanced production of ligands, overproduction of receptors and mutations leading to constitutive activation of receptors.¹⁰ We will shift our focus specifically to EGFR, and the mechanisms for its oncogenic activation as illustrated in **Figure 2.5**, as it is the therapeutic target of importance for this body of research.

2.1.4.1 EGFR Overexpression

The aberrant activation of EGFR is often associated with EGFR overexpression. There are different mechanisms by which EGFR is overproduced within tumour cells and the role of the enhanced level of EGFR in tumorigenesis and correlation with poor prognosis in a variety of cancers has been well documented.¹⁴ Amplification of the EGFR gene leads to increased levels of EGFR and this amplification has been observed in NSCLC, breast carcinomas and glioblastoma multiforme (GBM).¹⁵ EGFR may also be overexpressed due to deregulation within the signal-processing layer, with EGFR transcription being directly activated by the binding of wild-type and mutant p53 proteins to specific promoters.¹⁶

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

Tumour cells typically contain an abundance of mutant p53 proteins, which results in continuous promotor binding and transcription activation, leading to overexpression of the receptor. With increased levels of EGFR on the cell surface, spontaneous receptor dimerization and constitutive activation has been documented even in the absence of ligand binding, as shown in **Figure 2.5.1**. These constitutively active receptors lead to continuous activation of downstream signalling cascades.

2.1.4.2 Overproduction of Ligands

The ligands EGF and TGF- α have frequently been found to be co-overexpressed with EGFR in various cancers. As these ligands act in an autocrine/paracrine manner, their overexpression in tandem with EGFR in cancer cells activates autocrine loops, shown in **Figure 2.5.2**.^{12, 17} Autocrine loops, where a cell produces the same ligands for which it has receptors, provides cancer cells with self-sustaining growth signals and results in hyperactivation and uncontrolled tumour growth. Studies in mouse models have shown that overexpression of these growth factors lead to the induction of tumour formation and a higher rate of tumour cell proliferation, and are frequently found in nearly all forms of human NSCLC.^{12, 17}

2.1.4.3 EGFR Cross-talk and Heterodimerization

The communication of EGFR with other cell-surface receptors plays a widely functional role in tumour development. This cross-talk takes place between other members of the ErbB family, other receptor tyrosine kinases (RTK), cell adhesion molecules and G-protein coupled receptors.¹² As mentioned previously, the EGFR-ErbB2 heterodimer, shown in **Figure 2.5.3**, is the most potent inducer of mitogenic signalling and cell transformation and is commonly co-overexpressed in brain, breast and ovarian cancer.¹² Furthermore, the importance of these heterodimers in cancer cells has been demonstrated by the administration of antibodies blocking specific receptor heterodimerization in breast and prostate cancer xenograft models, resulting in powerful anti-tumour effects.¹⁸ The cross-talk of EGFR with several cell adhesion molecules within the extracellular matrix (ECM), such as integrin, has been shown to induce tyrosine phosphorylation of EGFR, independent of ligand binding.¹⁹ This has been postulated to occur by association of macromolecular complexes with EGFR on the cell membrane (**Figure 2.5.3**) and involves various signalling and adaptor molecules.¹²

2.1.4.4 Defective EGFR downregulation

While the hyperactivation of EGFR signalling pathways and its strong associations with carcinogenesis is well established, the impaired deactivation of these pathways has also been linked to neoplastic cell transformation.²⁰ The main mechanism by which EGFR signalling is deactivated, termed receptor downregulation, involves the ligand-induced internalization of the receptor through endocytosis and subsequent degradation of this receptor in lysosomes, as shown in **Figure 2.5.4**.¹² This process is carried out by complex molecular machinery that utilises the small protein ubiquitin as a key regulator, facilitating proper endocytosis and degradation.²⁰ Oncogenic forms of EGFR escape downregulation due to the deregulation or lack of integral components of this process, such as ubiquitin ligase c-Cbl. By binding to phosphorylated sites in the EGFR intracellular domain, c-Cbl directs the activated receptor to the lysosome for degradation by mediating tagging with ubiquitin.²¹

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

EGFR mutants with deletions spanning the docking sites of c-Cbl lack this binding, and mutant forms of c-Cbl have been found to bind to EGFR but fail to ubiquitinate the receptor, resulting in downregulation being avoided and constant activation of the receptor.¹²

2.1.4.5 Activating EGFR Mutations

Mutations in the EGFR encoding gene that constitutively activate the receptor are recurrently found in human cancers. These mutations can be classified by the region that they occur and are divided into three main groups, namely: mutations within the extracellular domain, the intracellular domain and the intracellular kinase domain.¹⁵ Mutations within the extracellular domain are typically observed in GBM, with amplification of the EGFR gene being particularly prominent.¹⁵ The majority of these occur through deletions of specific exons, encoding for all or parts of the extracellular domain. These shortened genes give rise to truncated receptors, forced into the active extended form, resulting in ligand-independent constitutive activation. A well-documented example of this is the EGFRvIII mutation, shown in **Figure 2.5.5**, which results from an in-frame deletion of exons 2 – 7.¹² These genes encode for the whole of subdomain I and a large portion of subdomain II, giving rise to a receptor trapped in a partially activated state with most of its ligand-binding area missing. This is sufficient for oncogenic signalling, being constitutively phosphorylated and able to activate tumorigenic signalling pathways such as RAS-ERK and PI3K-ATK, but insufficient for receptor degradation as the truncated receptor is unable to be ubiquitinated, avoiding downregulation and prolonging oncogenic signalling.^{22, 23} Mutations within the intracellular domain primarily consist of large deletions and/or duplications of exons and as with extracellular mutations, are best described in GBM.¹² These include mutations affecting the juxtamembrane and its stabilization of the asymmetric kinase dimer, and mutations in the C-terminal tail region shifting the equilibrium of the kinase domain towards an active asymmetric dimer.^{9, 24}

The last group of EGFR mutations bring about small changes in the seven exons (18 – 24) that code for the intracellular kinase domain and are very frequently associated with NSCLC.²⁵ These consist of small point mutations and deletions centred around the ATP binding cleft of the tyrosine kinase domain. These mutations can increase kinase activity by reorganizing critical amino acids and stabilizing ATP binding affinity, as well as allowing for autophosphorylation.²⁶ Most importantly, mutations within this region may confer drug resistance, as one strategy to abrogate hyperactivation of EGFR and its oncogenic signalling is the inhibition of receptor kinase activity using tyrosine kinase inhibitors (TKI). These small molecule inhibitors target the intracellular kinase domain of EGFR and mutations within the kinase ATP binding site invariably induce a reduced affinity and acquired therapeutic resistance to these drugs.²⁷ The understanding and development of novel drug candidates to combat these mutant, drug-resistant forms of EGFR, embodies the basis of this research project. The origin and mechanism of protein phosphorylation within kinases, the characterization of the human kinome and an overview of kinase inhibition as a therapeutic strategy will be laid out in the next section. This will provide the foundation for discussing the successes and failures in the history of therapeutic targeting of EGFR and its present state.

2.2 Protein Kinases

2.2.1 Origins of Reversible Protein Phosphorylation

While its central and essential role is recognized today, the significance of reversible protein phosphorylation was slow to be appreciated. The foundation of its discovery lay in the collaboration of Edwin Krebs and Edmond Fischer in the early 1950s, who demonstrated that the interconversion of phosphorylase b into phosphorylase a occurred readily in the presence of magnesium adenosine triphosphate (Mg-ATP) and an enzyme that they coined phosphorylase kinase. This kinase was subsequently shown to catalyse the transfer of a γ -phosphoryl group of ATP to a specific serine residue on phosphorylase b, facilitating conversion into phosphorylase a.²⁸ Furthermore, the reversion of phosphorylase a to b was shown to be activated by the release of this γ -phosphoryl group by a phosphate-releasing enzyme known today as phosphatase 1.²⁸ These discoveries, and later work concerning reversible phosphorylation as a biological regulatory mechanism, led to the duo being awarded the 1992 Nobel Prize in Physiology or Medicine.^{29, 30} The discovery of protein kinase A (PKA), and its roles in the activation of phosphorylase kinase and the inhibition of glycogen synthase, were the first examples of a kinase activation cascade and enzyme inhibition by phosphorylation respectively.³¹ At the end of the 1960s however, phosphorylation was regarded as a specialised regulatory mechanism confined to glycogen metabolism.

It was through discoveries in the 1970s and early 1980s that the significance of protein phosphorylation began to be acknowledged. Research undertaken by Lester Reed in 1969 revealed the inactivation of the mitochondrial pyruvate dehydrogenase complex by phosphorylation, alluding to its operation as a control mechanism in other metabolic pathways and organelles.³² The first example of multisite phosphorylation, where multiple residues of a protein are phosphorylated by multiple kinases, was witnessed in the 1970s.³³ Geneticists were granted a deeper functional understanding of numerous regulatory genes through the unravelling of the first protein kinase A (PKA) amino-acid sequence.³⁴ Finally, the landmark discovery by Ray Erikson that v-Src - the protein encoded by the Rous sarcoma virus transforming gene - was indeed a kinase, led Tony Hunter to the revelation that v-Src phosphorylates tyrosine residues in proteins, the first example of a protein tyrosine kinase.^{35, 36} In the following years, EGFR and many other growth factor receptors were shown to have protein tyrosine kinase domains, activated when the natural substrate engages the receptor in the extracellular membrane as discussed in **Section 2.1.2**.³⁷ These discoveries sparked an investigation into the role of protein kinases and phosphorylation in signal transduction pathways.

The 1990s has been dubbed by some as the decade of protein kinase cascades.³³ Following the identification of MAPK in the late 1980s, it was found to be activated by the phosphorylation of a threonine and tyrosine residue, catalysed by another “dual specificity” MAPK kinase, through a RAS dependant signalling pathway.³⁸ This classical MAP kinase cascade was elucidated in the early 1990s through the collaborative efforts of several laboratories and followed by the dissection of many other MAP kinase cascades, including those that are activated in the ErbB signalling network.³⁹ This in turn led to the discovery of other kinases and kinase cascades in the decade, including the PI3K-AKT-mTOR pathway. Since the discoveries of Fischer and Krebs nearly 70 years ago, our knowledge of reversible protein phosphorylation has dramatically increased, leading to the unravelling of the human kinome and identification of all human kinases and a deeper understanding of its mechanistic insights and regulatory functions.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

2.2.2 The Human Kinome

With the completion of the human genome sequence in 2001, the door was opened for the cataloguing of the protein kinase complement of the human genome. The human kinome consists of 518 putative protein kinase genes, making it one of the largest gene families.⁴⁰ This is a strikingly large number, making up nearly 2% of the entire human genome. The enzymes encoded by these genes catalyse the phosphorylation of various OH groups present on proteins and are classified based upon the nature of this phosphorylated OH group. These are protein-serine/threonine kinases (385 members), protein-tyrosine kinases (90 members) and tyrosine-kinase like proteins (43 members).⁴⁰ Of the 90 protein-tyrosine kinases, 58 are receptor kinases (of which EGFR is one) and 32 are non-receptor kinases. A small group of dual-specificity kinases possess the ability to catalyse phosphorylation of both tyrosine and threonine residues in target proteins, such as MEK1 and MEK2 involved in the MAP pathways discussed in the previous section. These kinases are placed within the protein-serine/threonine kinase family due to molecular features. Protein kinases play a predominant regulatory role in cell differentiation and growth, development of the immune system and nervous system function, transcription and apoptosis.² Consequently, the dysregulation of these enzymes has been connected to a variety of diseases such as cancer, cardiovascular, autoimmune, inflammatory and nervous diseases and diabetes.⁴¹ The understanding of the physiological and pathological functions of protein kinases and their underlying mechanisms has thus become a target of intense interest in the last 30 years.

2.2.3 The Mechanism of Protein Phosphorylation and the Kinase Active Site

The reversible phosphorylation of proteins regulates nearly every aspect of cellular life and is the most prevalent method of post-translational modification used in signal transduction pathways.^{33, 42} This process has a regulatory role in cellular metabolism, division, growth, motility, differentiation, membrane transport, muscle contraction, immunity, learning and memory.⁴² It is estimated that 30% of all cellular proteins are phosphorylated on at least one residue, equating to a typical kinase having to differentiate between approximately 700 000 potential phosphorylation sites.⁴² This exquisite specificity results from the contribution of multiple mechanisms from the kinase catalytic site, interaction between kinase and substrate and error correcting mechanisms.

Protein phosphorylation is a process in which a protein kinase, shown in brown in **Figure 2.6a**, catalyses the transfer of a γ -phosphate group from ATP to a specific amino acid on a protein substrate (**Figure 2.6a**, yellow). In eukaryotes these substrate amino acids are usually serine, threonine or tyrosine residues and after this transfer ADP is liberated. Protein phosphatases catalyse the dephosphorylation of protein substrates, making this an overall reversible process (**Figure 2.6a**). Protein kinases share a basic catalytic cycle for phosphorylation, starting with the binding of ATP to the active site of the kinase as seen in the top left of **Figure 2.6b**. This is followed by binding of the substrate (**Figure 2.6b**, green) to the active site, allowing for γ -phosphoryl (**Figure 2.6b**, red) transfer to a specific residue on the substrate. Following phosphorylation, the substrate and lastly ADP are released from the binding site of the kinase.⁴² There is variation in the order of the steps of the catalytic cycle amongst different kinases. For instance, substrates may bind to the kinase active site prior to ATP binding and ADP may be released before the substrate following phosphorylation. The rate limiting step of this process may also vary between different kinases.⁴²

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

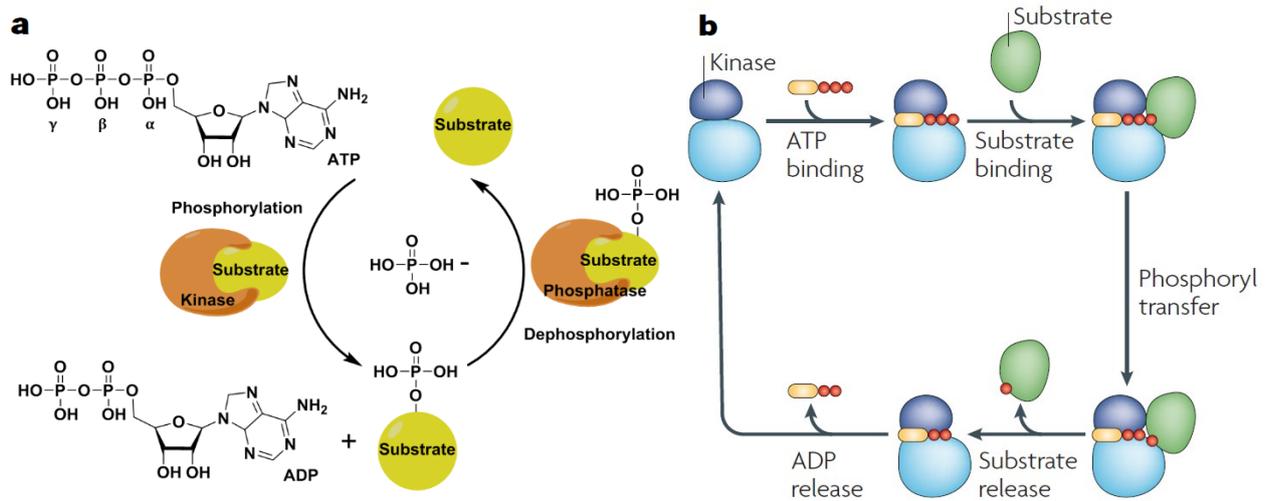


Figure 2.6: a) The reversible phosphorylation of protein substrates by kinases and phosphatases and b) The basic catalytic cycle for substrate phosphorylation by a kinase. Figure reproduced from b) Ubersax et al.⁴²

Most eukaryotic protein kinases are structurally similar, with a catalytic domain of roughly 250 amino acids in length. The canonical protein fold consists of two lobes, a smaller N-lobe comprising of five antiparallel β -strands and one helix and a larger C-lobe consisting of α -helices, illustrated in **Figure 2.7a** with EGFR as example.⁴³ The ATP-binding cleft is a narrow hydrophobic pocket found between the N- and C-terminus of these two lobes, linked by a flexible hinge region. The hinge region, shown in **Figure 2.7b**, typically has one hydrogen bond donor flanked by two hydrogen bond acceptors, derived from the protein backbone. ATP binds in a manner that facilitates the adenosine moiety (**Figure 2.7b**, dark grey) forming important reversible interactions with the hinge region residues and so that the hydrophilic phosphate backbone (**Figure 2.7b**, light grey) is oriented towards, and exposed to, the solvent.⁴⁴ Reversible hinge region interactions with certain kinase inhibitors is critical to their efficacy and is a primary consideration in the design of potential drug candidates. The P-loop, or the glycine rich loop, forms the roof of the active site, with C-terminal β -sheets covering the bottom. The important gatekeeper residues control access to the specificity pocket (**Figure 2.7b**, yellow).

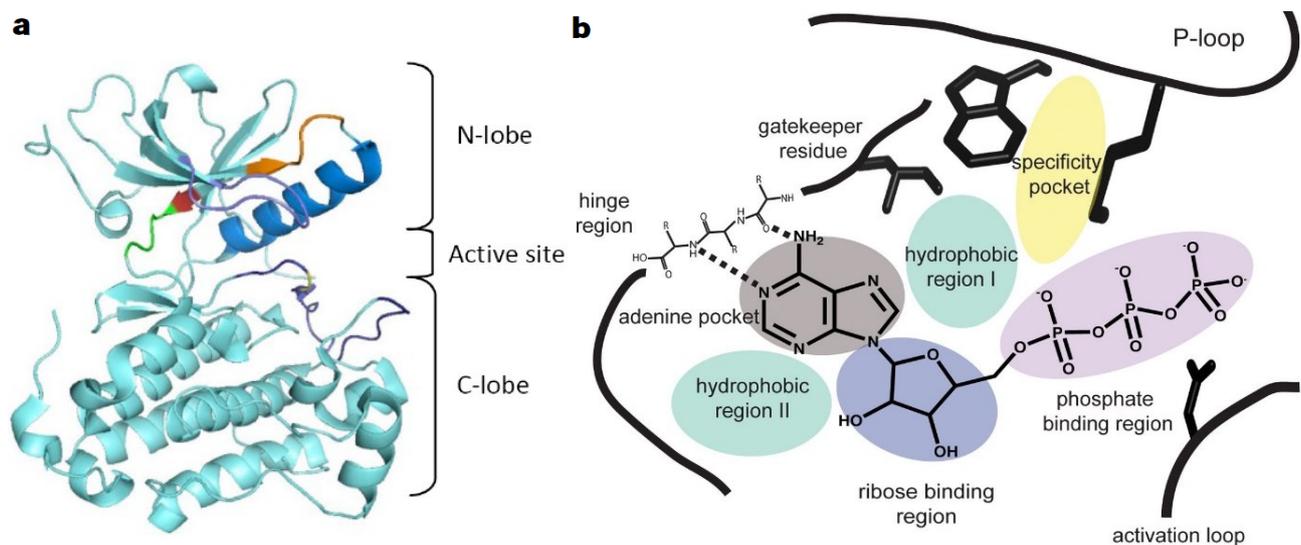


Figure 2.7: a) Protein ribbon view of EGFR showing the N- and C-lobes with the active site between (PDB: 2GS2) and b) A typical protein kinase ATP binding pocket. Figure reproduced from a) Warnault et al and b) Williams et al.^{43, 45}

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

All protein kinases have a flexible activation loop (**Figure 2.7b**), starting with the conserved amino acid sequence Asp-Phe-Gly (DFG), which controls access to the active site and contains the serine, threonine and tyrosine residues which may be phosphorylated. When these residues are not phosphorylated, this loop occupies a large area of the active site, effectively blocking it. Upon activating-phosphorylation by another upstream kinase (or through asymmetric dimer formation in EGFR), the hydrophilic, charged activation loop unfolds and moves into the solvent to expose the ATP binding pocket. This enables ATP binding and transfer of the phosphate group to a substrate protein or downstream kinase.⁴⁴ The conformation of the DFG motif plays an important role in the type of kinase inhibitor binding conformation and will be discussed in the following section.

Human protein kinases share a high degree of similarity in their 3D structure, which is most pronounced in the kinase active site. However, structural features such as the difference in the charge distribution and hydrophobicity of their surface residues, and a deeper catalytic cleft found in protein tyrosine-specific kinases compared to that of serine/threonine-specific kinases, allow for substrate phosphorylation specificity.^{42,46} Within the ATP binding pocket, important residues such as the gatekeeper side chain, dictates access to the hydrophobic regions (**Figure 2.7b**, light blue) and specificity pocket (**Figure 2.7b**, yellow) located within the active site. A deeper understanding of these structural subtleties has allowed for the informed design of kinase inhibitors and illuminated various mechanisms of inhibition. Over the past two decades, drug design strategies stimulated by these factors have given rise to one of the most intensively pursued and successful classes of targeted therapeutic agents – protein kinase inhibitors. We will examine this diverse class of inhibitors and its varied approaches in the following section.

2.3 Protein Kinase Inhibitors

2.3.1 Overview

Many of the 518 protein kinases encoded by the human genome have been established as prominent therapeutic targets due to deregulation of their function in signal transduction networks and its implication in many diseases. This class of enzyme is the second most targeted group in drug discovery after the G-protein coupled receptors.⁴⁷ Since the 2001 milestone approval of the first kinase inhibitor for human use, Imatinib, 37 other drug candidates have won FDA approval to date.⁴⁸ Strikingly, half of these approvals have occurred within the past 4 years, signifying the continued relevance of protein kinases as a therapeutic target. It is estimated that approximately 250 kinase inhibitors are undergoing clinical phase trials, with a database of nearly 220 000 papers in PubMed and more than 47 000 patents and patent applications found in SciFinder on kinase inhibitors.⁴⁸ Over 5000 crystal structures have been solved with or without small molecules, inhibition assays have been developed for more than 80% of the human kinome and small molecule kinase inhibitors have been developed to target nearly 20% of the human kinome.⁴⁹ Kinase inhibitors are classified based upon the binding mode of the inhibitor and the unique conformational form of the kinase targeted. Kinase inhibitors can be ATP or non-ATP competitive, reversible or irreversible, allosteric, bi-valent or a combination of these features. This determines the category of the kinase inhibitor, which ranges from type I-VI. The different types and their properties will be explored in the following section.

2.3.2 Types of Kinase Inhibitors

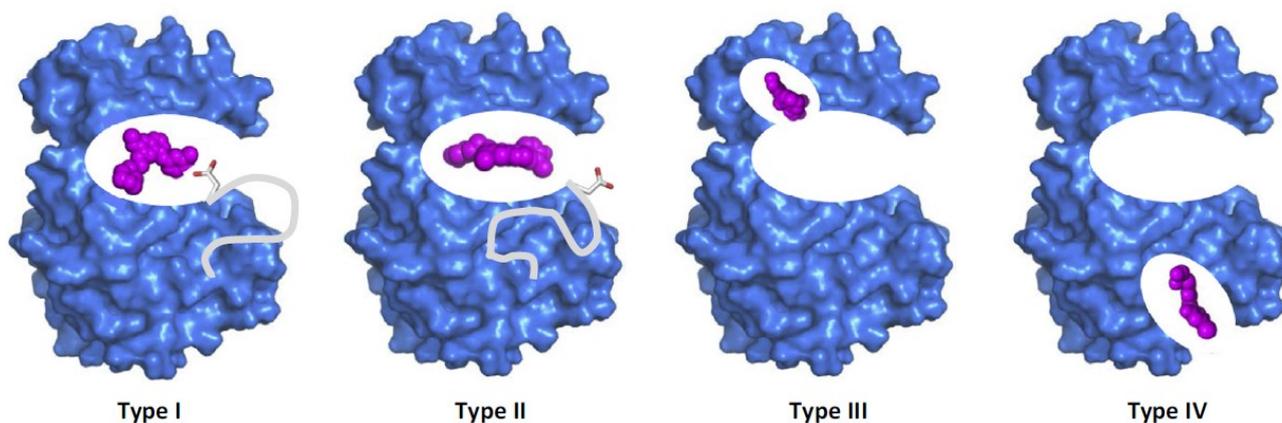


Figure 2.8: The four types of reversible binding modes for kinase inhibitors. Figure reproduced from Wu et al.⁴⁹

2.3.2.1 Type I and II Inhibitors

Type I and II represent ATP-competitive inhibitors that mimic the purine ring of the adenine moiety of ATP, forming hydrogen bond interactions with one or more hinge residues. Their classification depends on the activation state of the protein kinase, or more specifically the disposition of the DFG-Asp motif, being either active in or inactive out. Type I inhibitors, as seen on the far left of **Figure 2.8**, bind to the phosphorylated active conformation of the kinase in a reversible manner, with the aspartate residue (**Figure 2.8**, white backbone) of the DFG motif facing into the ATP-binding pocket of the kinase (DFG-in).⁴⁹ These inhibitors usually contain a heterocyclic ring system, or driving group, that occupies the purine binding site. This serves as a scaffold for various side chains to occupy and interact with adjacent hydrophobic regions, including the gatekeeper residue. Solubilising groups are typically employed in the hydrophilic regions of the enzyme usually occupied by the ribose moiety of ATP.⁴⁷ Several type I inhibitors have received FDA approval for the treatment of cancer, including the EGFR inhibitors erlotinib, gefitinib and lapatinib.⁵⁰ Due to the conserved nature of the ATP binding pocket throughout the human kinome, type I inhibitors display an inclination for low selectivity, increasing the potential for off-target and adverse side effects.

Type II inhibitors, (**Figure 2.8**, second from left) act by targeting and stabilizing the unphosphorylated inactive conformation of the kinase, with the now flipped aspartate residue (**Figure 2.8**, white backbone) of the DFG motif protruding outward from the ATP-binding site (DFG-out).⁴⁹ In addition to reversible contacts with the hinge region, type II inhibitors exploit accessibility to, and new interactions within, specific lipophilic pockets adjacent to the ATP-binding site, derived from the conformational change of the DFG motif. This enhances their selectivity profile between kinases, but overlap between type I and II inhibitors is still considerable.⁴⁷ FDA approved type II inhibitors include imatinib, nilotinib and ponatinib used in the treatment of CML.⁵⁰

2.3.2.2 Type III and IV Inhibitors

Type III and IV inhibitors bind outside of the catalytic domain/ATP-binding site and modulate kinase activity in an allosteric manner, as both ATP and the allosteric inhibitor can bind simultaneously to the kinase.⁴⁷ Type III inhibitors, (**Figure 2.8**, second from right) bind in an allosteric pocket adjacent to the ATP-binding site without interacting with it, whilst type IV inhibitors (**Figure 2.8**, right) occupy an allosteric pocket distant from the substrate-binding sites.⁴⁹

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

Type III and IV inhibitors are ATP non-competitive, as the binding of ATP does not prevent their interaction with the targeted kinase. Allosteric inhibitors show the highest degree of selectivity in targeted enzyme inhibition, as they take advantage of binding sites and physiological mechanisms that are exclusive to particular kinases.⁴⁷ Owing to these attractive characteristics, the use of allosteric inhibitors is thought to be critical in overcoming selectivity, off-target effects and drug resistance.⁴⁷ Trametinib is currently the only type III FDA approved drug, used for the inhibition of the MAPK enzymes MEK1 and MEK2.⁵⁰

2.3.2.3 Type V Inhibitors

Type V inhibitors have been labelled as bivalent molecules that span two regions of the protein kinase domain.⁵¹ As such, these bi-substrate or bivalent inhibitors exhibit more than one binding mode. Bi-substrate inhibitors target the ATP and other protein substrate binding sites, whilst bivalent inhibitors comprise of a ATP binding site moiety and a non-substrate ligand, targeting the ATP binding cleft and any surface except that of a protein substrate binding site.⁵¹ These inhibitors display the high efficacy and potency associated with type I and II inhibitors, whilst retaining the selectivity of allosteric inhibitors.

2.3.2.4 Type VI Inhibitors

The final class, type VI inhibitors, bind covalently to their protein kinase target and are known as covalent or irreversible inhibitors. The chemical rationale of these inhibitors is based on the presence of an exposed nucleophilic cysteine residue, in or around the ATP binding site of a small number of kinases, which is targeted for covalent modification with a suitable electrophile.⁵² This modification occurs via trapping of the cysteine residue either by an SN2 displacement of a leaving group or by a Michael addition reaction with a suitable Michael acceptor incorporated in the inhibitor. As illustrated in **Figure 2.9a**, the α -, β -unsaturated carbonyl undergoes a 1,4-conjugate addition with the nucleophilic cysteine, followed by proton abstraction, to form a covalent bond with the cysteine residue. The formation of the Michael adduct is a reversible process and the mechanism of the cysteine-involved Michael Addition reaction, with microenvironmental effects, has been widely studied within the scientific community.⁵³

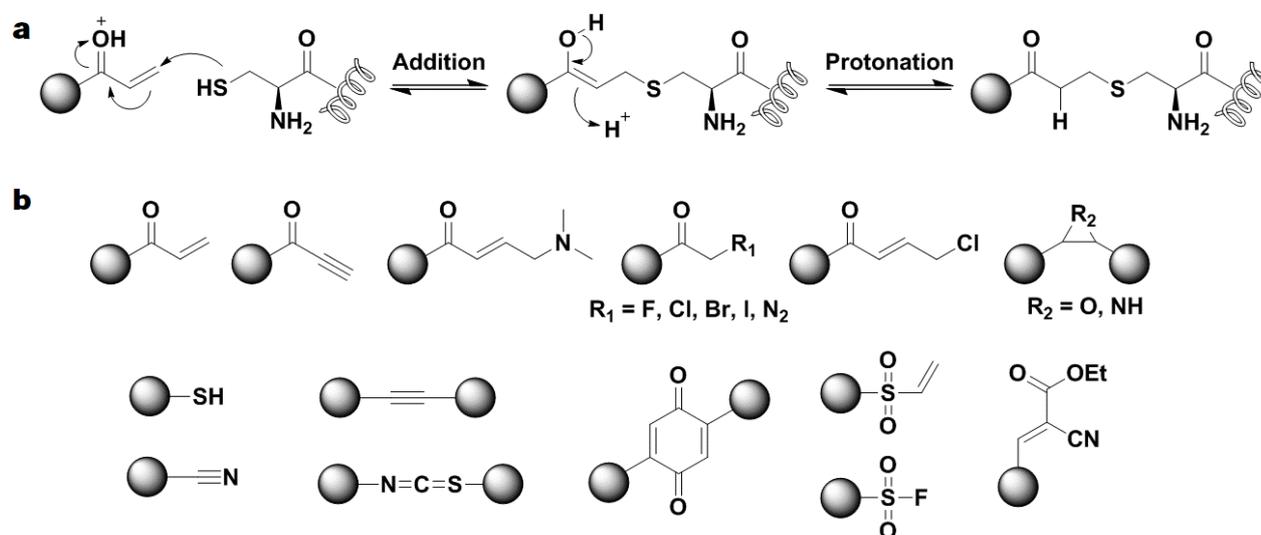


Figure 2.9: a) Mechanism illustrating the Michael Addition reaction of cysteine to an acrylyc electrophile and b) Electrophilic "warheads" used in irreversible inhibitors.⁵²

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

It is broadly accepted that high affinity, reversible interactions within the binding pocket are primarily important, with the inhibitor initially binding non-covalently. This is followed by positioning of the reactive moiety into a suitable orientation that allows for irreversible enzyme modification.⁵² This underlines the importance of the correct positioning of the electrophile on the inhibitor scaffold, as well as the type of electrophile used. Most covalent inhibition strategies utilised to date target the nucleophilic cysteine thiolate and employ the use of a weakly reactive acrylamide moiety. Several other electrophilic “warheads”, shown in **Figure 2.9b**, which can react with cysteine and other nucleophiles such as lysine, tyrosine and serine are being increasingly explored. These include the Michael addition-specific vinyl sulfonates, quinones, alkyl amides and propargylic acid derivatives. Warheads such as α -halo ketones, thiocyanates, alkynes, nitriles, epoxides, sulfonyl fluorides or thiol itself which can undergo nucleophilic displacement or addition reactions.⁵² Lastly, in recent years the use of a cyano-acrylamide electrophile as a reversible covalent inhibitor has shown prolonged on-target residence times and superior efficacy.⁵⁴

The use of irreversible inhibition strategies holds many advantages over traditional reversible inhibitors, the first being afforded by targeting the potential protein kinase cysteinome. To date, 211 kinases have been identified as bearing approachable cysteine residues, while only six distinct sites have been targeted by a covalent inhibitor.⁵⁵ The finite number of candidates with a druggable cysteine residue enhances selectivity profiles of potential inhibitors. As irreversible inhibitors are ATP non-competitive, they are unaffected by changes in ATP binding affinity brought about by mutations and do not have to compete with high concentrations of endogenous ATP. As these inhibitors form a covalent bond, they display a prolonged residence time and extended dissociation half-life, improved pharmacodynamics and greater potency.⁴⁷ Historically, the risk of haptization, recurrent toxicity profiles and promiscuity associated with irreversible inhibitors, has meant a reluctance to pursue this class of inhibitors.⁵² However, in recent years, successes displayed by the FDA approved EGFR inhibitors afatinib and ibrutinib has sparked a resurgence in the development of covalent inhibition strategies.⁵⁰

The interaction of each inhibitor with a protein kinase target is unique. The classification of these interactions, and its application in the drug discovery process, has been invaluable, resulting in the FDA approving 2 – 4 small molecule kinase inhibitors per year.⁵⁰ Owing to its clinical relevance, multiple strategies have been developed to overcome the inevitable emergence of resistance mechanisms. This is no more pronounced than in the colourful history of EGFR inhibition. A timeline spanning two decades, marred with the onset of numerous mutations and three generations of blockbuster drugs to combat it, tyrosine kinase inhibitors are now standard in the treatment of NSCLC, breast cancer and a host of other diseases. Next, we will delve into the history of EGFR as a therapeutic target, looking at successes and failures, the onset of mutations in patients and the strategies used to overcome these hurdles.

2.4 EGFR as Therapeutic Target

The two distinct therapeutic methods that have been employed in the targeted inhibition of EGFR are the use of monoclonal antibodies and protein tyrosine kinase inhibitors (TKI). Each of these approaches have a unique mechanism of action, with EGFR antibodies binding to the extracellular domain, and TKIs to the intracellular kinase domain. As it is the most pertinent to our research theme, we will focus solely on the history and development of FDA approved TKIs in the next section, which will be described in terms of inhibitor “generations.”

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

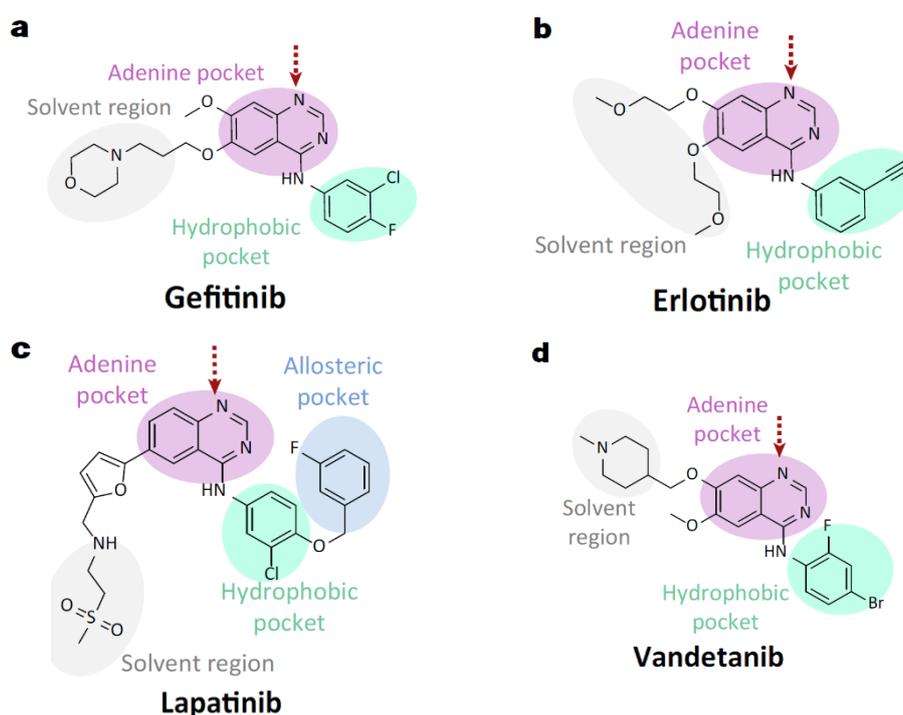
2.4.1 1st Generation Inhibitors

Figure 2.10: 1st generation EGFR inhibitors that have received FDA approval with red arrows denoting atoms forming reversible interactions with the hinge region. Figure reproduced from Wu et al.⁴⁹

In 2003, twenty years after it was proposed that therapeutic agents used to disrupt EGFR signalling could potentially treat cancer, gefitinib became the first TKI to receive FDA approval for treatment of NSCLC.⁵⁶ The anilinoquinazoline-derived inhibitor, shown in **Figure 2.10a**, was first characterized in 1996 and is prescribed to patients after failure of platinum-based or docetaxel chemotherapy.⁵⁷ Gefitinib is a type I inhibitor which binds in the ATP pocket, with the 4-anilinoquinazoline moiety forming reversible interactions with the hinge region (**Figure 2.10a**, red arrow). The drug shows a 200-fold greater affinity for EGFR in comparison to other ErbB family members and is orally active. A year later, another type I anilinoquinazoline-derived inhibitor, erlotinib (**Figure 2.10b**), received FDA approval for NSCLC.⁵⁸ Similar to gefitinib, erlotinib functions as an ATP competitive inhibitor, approved for patients with relapsed NSCLC and maintenance therapy in advanced NSCLC with no progression after four cycles of first-line chemotherapy. These two drugs are considered the 1st generation inhibitors of EGFR. Lapatinib (**Figure 2.10c**), was the first dual inhibitor of EGFR and HER2 to receive FDA approval in 2007.⁵⁹ It has also been found to exhibit inhibitory activity against AKT overexpression in human tumour xenografts. Due to its nonspecific nature of inhibition, it accounts for a broader spectrum of antitumor activity with higher efficacy and has shown significant inhibitory effects on human breast cancer cells.⁵⁷ Lastly, vandetanib (**Figure 2.10d**) acts as an inhibitor of a number of cell receptors, including rearranged during transfection (RET) receptor protein tyrosine kinase, VEGFR and EGFR. Vandetanib received FDA approval in 2011 for the treatment of thyroid cancer.⁶⁰

In 2004, less than a year after receiving its FDA approval and in response to treatment with gefitinib, two independent research groups simultaneously identified the presence of somatic EGFR mutations in a sub-group of NSCLC patients.^{26, 61}

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

The somatic mutations discovered chiefly targeted exons 18 – 21 of the EGFR gene, encoding for part of the tyrosine kinase domain around the ATP binding pocket of EGFR. The most frequent and well-documented of these mutations include in-frame deletions in exon 19, known as del19, eliminating the conserved residues 747 – 750 and the most famous exon 21 L858R substitution, which results in an amino acid substitution at position 858 from a leucine (L) to an arginine (R).⁶² These mutations make up between ~80 – 90% of all EGFR mutations in NSCLC and are referred to as activating mutations, altering the receptor into a constitutively activated oncogenic state which disrupts autoinhibitory interactions and accelerates catalysis by as much as 50 fold.^{63, 64}

Interestingly, these mutant enzymes showed reduced binding affinity towards ATP and an increased sensitivity towards EGFR TKIs, such as gefitinib and erlotinib, as they bind in a reversible and ATP-competitive manner.⁶⁴ Gefitinib has also been shown to have a 20-fold increased binding affinity towards the L858R mutant than to the wild-type enzyme.⁶⁴ Serendipitously, these two effects combine to yield the remarkable potency of gefitinib and erlotinib against tumours that exhibit an “oncogenic addiction” to the mutant activated form of EGFR for survival.⁶⁵ This afforded the excellent preliminary clinical activity and *in vitro* sensitivity of gefitinib, leading directly to its FDA approval.⁶³ While this demonstrates that certain mutations, while oncogenic in nature, can be exploited and benefitted from in targeted therapy, other mutations may result in the opposite and confer resistance.

2.4.2 2nd Generation Inhibitors

While the use of gefitinib and erlotinib as first line therapy for NSCLC patients with activating mutations was lauded as a scientific achievement, the success was short-lived. Nearly all patients who initially benefitted from this treatment eventually acquired resistance through the emergence of various mechanisms. The most notorious of these, the secondary mutation known as T790M, was uncovered by two independent groups in 2005.^{27, 66} The EGFR T790M mutation is the primary mechanism of acquired resistance in tumours of patients undergoing EGFR TKI therapy, accounting for ~50 – 60% of all cases.⁶⁷ The mutation occurs in exon 20, with substitution of the threonine 790 amino acid, shown in **Figure 2.11a**, with a bulky methionine (**Figure 2.11b**), located within the ATP binding pocket.⁴⁹ Threonine 790 is the gatekeeper residue in EGFR, making it an important determinant of inhibitor specificity in protein kinases.⁶⁸ The mutation resulted in the dramatic loss of binding affinity of 1st generation inhibitors, rendering them ineffectual. Owing to the bulky nature of the molecular residue and key location at the entrance of the hydrophobic pocket within the ATP binding cleft, resistance was initially thought to be caused by steric interference of the methionine residue with 1st generation inhibitors.²⁷ This clash can be seen when comparing the binding mode of gefitinib to EGFR wild-type in **Figure 2.11a**, with that of the EGFR L858R/T790M double mutant shown in **Figure 2.11b**, altering the orientation and interactions of the inhibitor.⁴⁹

However, further studies indicated that the primary mechanism of resistance was mediated through an increase in protein binding affinity for ATP.⁶⁸ The T790M mutation, in conjunction with activating mutations, was shown to lower the concentration of ATP required to achieve a half-maximal reaction rate. These biochemical changes returned the ATP binding affinity of the double mutant to a level similar to wild-type EGFR, nullifying the advantages of the ATP-competitive 1st generation inhibitors and rendering them ineffective.⁶⁸

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

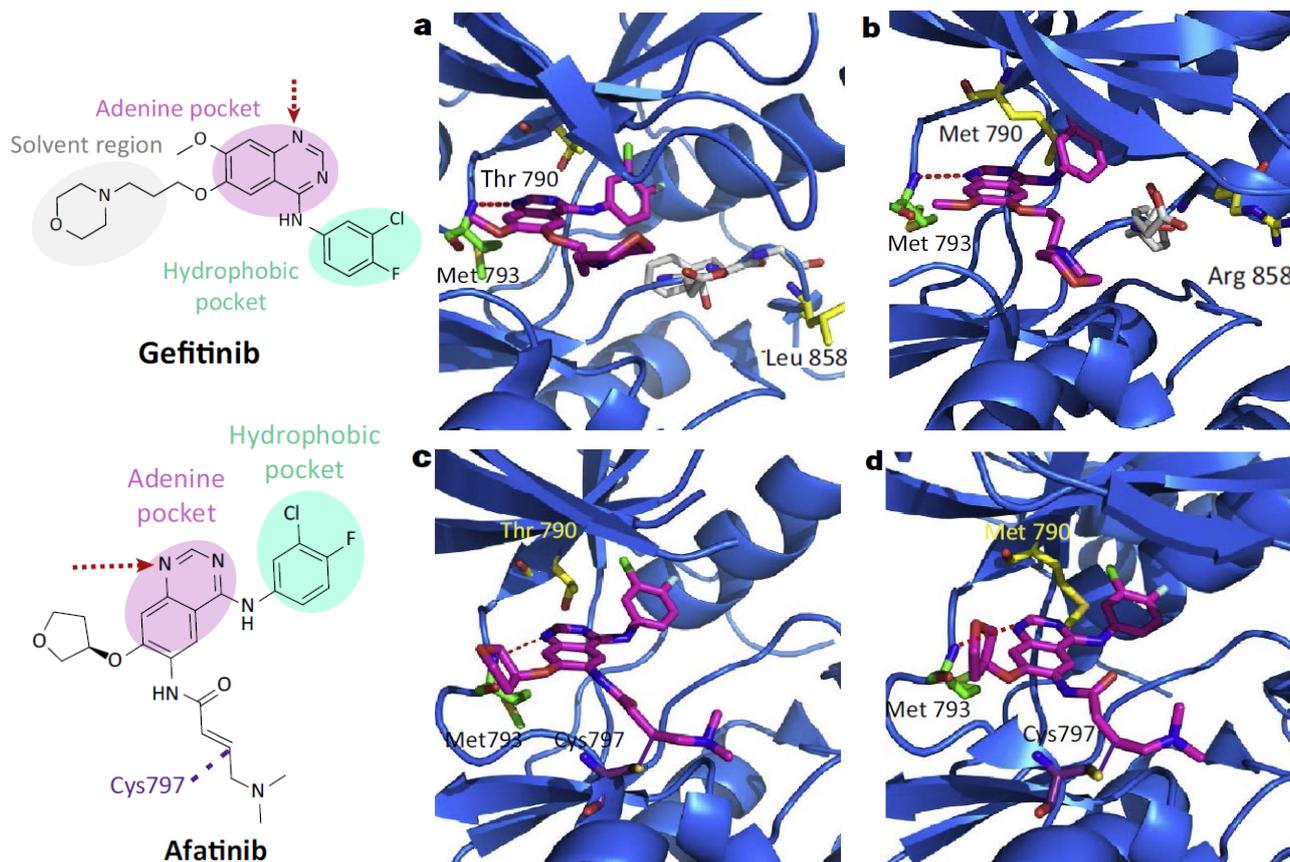


Figure 2.11: Binding modes of gefitinib with the a) EGFR wild type (PDB: 2ITY), b) EGFR L858R/T790M double mutant (PDB: 4I22) and afatinib with the c) EGFR wild type (PDB: 4G5J), d) EGFR L858R/T790M double mutant (PDB: 4G5P).

Figure reproduced from Wu et al.⁴⁹

Furthermore, studies showing that the single mutant T790M EGFR conferred a growth disadvantage in comparison to EGFR with activating mutations - with decreased ATP throughput and a lower enzymatic turnover - further supported this reasoning.⁶⁹

The onset of resistance mutations necessitated a paradigm shift and new strategy for targeted inhibitors of EGFR. Researchers turned their attention towards the development of irreversible inhibitors, specifically targeting covalent modification of cysteine residues within the ATP catalytic site with appropriate warheads, as discussed in **Section 2.3.2.4**. Following the identification of the T790M mutation, a wealth of potential drug candidates was developed, collectively forming the 2nd generation of EGFR inhibitors. While many of these entered late stage phase III clinical trials, almost all ended in limited success or discontinuation. One inhibitor developed within this period, afatinib, pictured above in **Figure 2.11**, was shown to confer a progression free survival (PFS) advantage over systematic chemotherapy and gefitinib as a first line treatment for NSCLC.^{70, 71} While structurally similar to gefitinib, afatinib binds irreversibly to cysteine 797 with a dimethylamino-acrylamide warhead (**Figure 2.11**). Tumours with activating mutations, such as L858R, were found to be particularly responsive to this 2nd generation inhibitor, showing longer overall survival in patients when compared to systematic chemotherapy.⁷¹ For this reason, afatinib was the only 2nd generation inhibitor to win FDA approval in 2013 as a first line treatment for metastatic EGFR-mutated NSCLC.⁷²

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

In preclinical models afatinib was shown to inhibit the T790M mutant receptor.⁷³ However, in clinical use as salvage therapy after 1st generation inhibitors, afatinib was unable to effectively overcome resistance mutations at clinically achievable doses.⁷⁴ The T790M mutation has been shown to be the main mechanism of resistance to afatinib in patients, originating from a combination of the steric interactions shown in comparison of **Figure 2.11c** and **Figure 2.11d**, as well as the enhanced binding affinity towards ATP discussed previously.

While not represented visually, it is worthy to note that neither dacomitinib nor neratinib, another two 2nd generation inhibitors, were able to demonstrate significant activity as a single agent against patients harbouring the T790M resistance mutation.^{75, 76} Amidst negative reviews due to extreme side effects, neratinib has since received FDA approval as of July 2017 as adjuvant treatment for HER2 overexpression/amplified breast cancer, while an application for FDA approval was filed in August 2018 for dacomitinib. To overcome the limitations and failures experienced with the development of 2nd generation inhibitors, researchers came to the realization that they would have to target the T790M mutation specifically. This led to the emergence of a novel class of mutant selective inhibitors, known as 3rd generation inhibitors.

2.4.3 3rd Generation Inhibitors

The lack of success, coupled with toxicity due to concurrent inhibition of wild-type EGFR associated with quinazoline-based 2nd generation EGFR inhibitors, left an urgent need for a novel therapeutic treatment for NSCLC patients who suffered from T790M drug resistance. Researchers reassessed the development of EGFR inhibitors, with the EGFR-T790M receptor being considered as an entirely new enzyme and not a mutant form of the wild-type. The following features were identified as critical to overcome and to achieve clinical efficacy against T790M drug resistance: i) a novel scaffold core to specifically target the T790M gatekeeper mutant variant by avoiding the Met790 steric clash, ii) sparing wild-type inhibition by being mutant selective and iii) forming a covalent bond with Cys797 to increase drug-target residence time and circumvent ATP competition.⁷⁷ By applying these considerations, a covalent aminopyrimidine-based EGFR inhibitor, **WZ4002**, was identified in 2009 and is shown in **Figure 2.12a**.⁷⁸

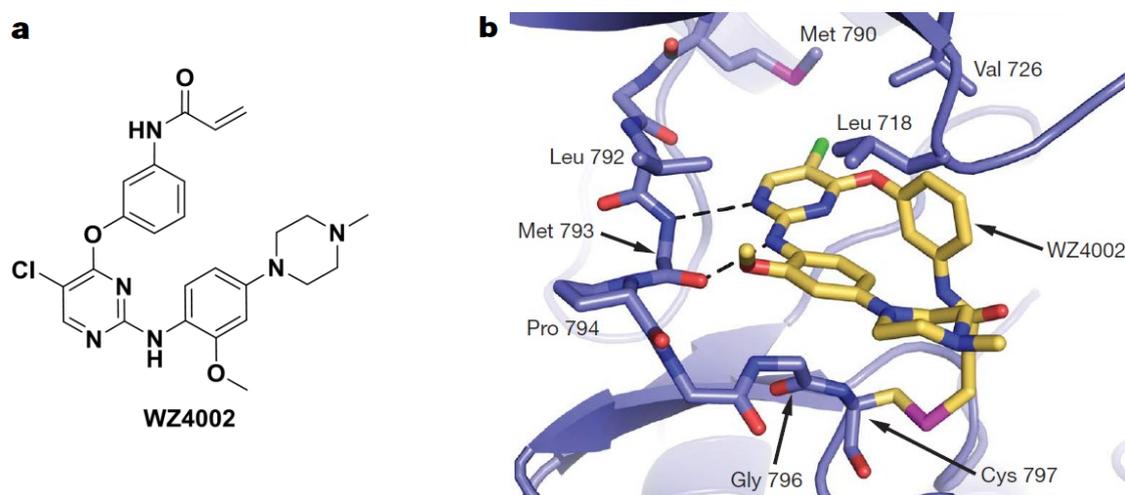


Figure 2.12: a) Structure of WZ4002 and b) Crystal structure illustrating binding mode and covalent inhibition (PDB: 3IKA).

Figure reproduced from b) Zhou et al.⁷⁸

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

WZ4002 is considered the prototype 3rd generation inhibitor, displaying a 100-fold increase in potency against EGFR-T790M and being up to 100-fold less potent against wild-type EGFR *in vitro* when compared to the quinazoline-based 1st and 2nd generation inhibitors.⁷⁸ WZ4002 was also shown to be effective in murine models of lung cancer driven by the T790M mutation. A crystal structure of the binding mode, shown in **Figure 2.12b**, revealed the structural basis for the increased potency and mutant selectivity, as WZ4002 was able to avoid the steric conflict with Met790 and form a covalent bond with Cys797.

While WZ4002 did not progress into human clinical trials, the marked success and ability of the first in class inhibitor to overcome T790M mutation resistance sparked the development and synthesis of various aminopyrimidine-based structural analogues including **rociletinib** and **olmutinib**, shown in **Figure 2.13** below. Rociletinib was developed in 2013, retaining the efficacy and mutant selective characteristics of WZ4002, with patients displaying impressive response rates and a median PFS of 13 months.⁷⁹ Further development of rociletinib was halted due to adverse events of hyperglycaemia in patients, caused by a metabolite of the candidate binding and inhibiting insulin growth factor receptor 1 (IGFR1).⁸⁰ Phase II clinical trials conducted with olmutinib delivered promising results, with little to no side effects, leading to a breakthrough therapy designation from the FDA in 2015 and approval for second line treatment in South Korea in 2016.⁸¹ However, amidst several controversies in the years following this approval, including the death of a patient from toxic epidermal necrolysis, the program was no longer considered commercially viable and terminated in early 2018.

The gold standard in EGFR TKI treatment of NSCLC is undoubtedly **osimertinib** (**Figure 2.13**). The drug discovery program that began in 2009, and led to the discovery of osimertinib, followed an adaptive approach and involved close collaboration with industrial partners and the global regulatory bodies of the FDA and European Medicines Agency (EMA).⁸² The project progressed with exceptional speed, delivering the clinical candidate in under 36 months. Following a clinical development of 2.5 years, osimertinib received breakthrough therapy designation in 2014 and provisional accelerated approval by the FDA in 2015.⁸² This was followed by provisional approval under an accelerated process by the EMA in the following year, the first under this program.⁸² Osimertinib became the first TKI to receive approval from both regulatory bodies in 2017, as first line treatment for patients with the EGFR-T790M mutation.

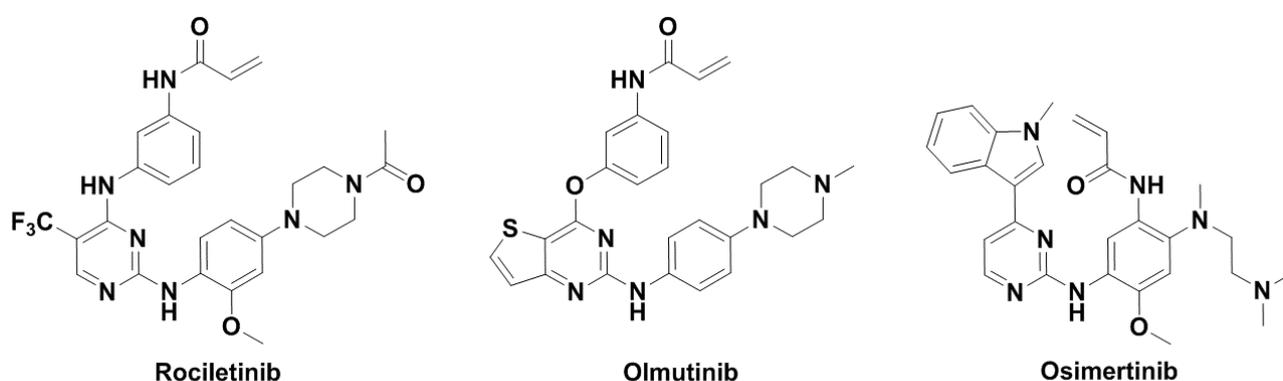


Figure 2.13: Chemical structures of 3rd generation inhibitors rociletinib, olmutinib and osimertinib.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

Osimertinib is architecturally unique, with a chemical structure distinct from other 3rd generation TKIs. When comparing the inhibitors depicted in **Figure 2.13**, these distinctions become apparent and include the electrophilic functionality residing on a different ring, an indole substituent on the pyrimidine ring and the use of N,N,N'-trimethylethylenediamine as solubilising group.⁸³ Unlike rociletinib, osimertinib and its resulting metabolites were specifically designed to show selectivity against IGFR1, minimizing potential dose limiting toxicities.⁸⁰ However, the unparalleled success of osimertinib for treatment of patients with the EGFR-T790M mutation has been overshadowed by the development of a common mechanism of acquired drug resistance toward covalent inhibitors.

Following treatment with 3rd generation inhibitors, including osimertinib, lung cancer patients developed resistance within a period of 10 months.⁸⁴ In 2015, researchers identified the emergence of a point mutation of the covalent anchor cysteine 797 to a less reactive serine residue, known as the EGFR-C797S mutation (**Figure 2.14a**). Irreversible inhibitors rely on the nucleophilicity of this cysteine residue to form a covalent bond with the receptor, as highlighted in yellow in **Figure 2.14b**, with WZ4002 as example. Following the tertiary C797S point mutation, formation of this covalent bond with currently available irreversible inhibitors is prevented (**Figure 2.14b**, highlighted in red), as the substituted hydroxyl of the serine residue is unlikely to undergo a Michael addition at physiological conditions, resulting in a considerable loss of efficacy.⁷⁷

Due to the selective pressure of 3rd generation inhibitors, specifically targeting the EGFR-T790M mutant and forming this covalent bond, it is not surprising that cancer cells acquire resistance through a tertiary mutation and the cysteine 797 position. This mechanism of resistance resembles the acquired C481S mutation observed in Bruton tyrosine kinase (BTK), which bears an analogous cysteine to EGFR, in response to the irreversible BTK inhibitor ibrutinib used for treatment of patients with chronic lymphocytic leukaemia (CLL).⁸⁵ This suggests that cysteine point mutations may constitute a common mechanism of acquired resistance and should therefore be considered a recurring liability to irreversible kinase inhibitors in the future.

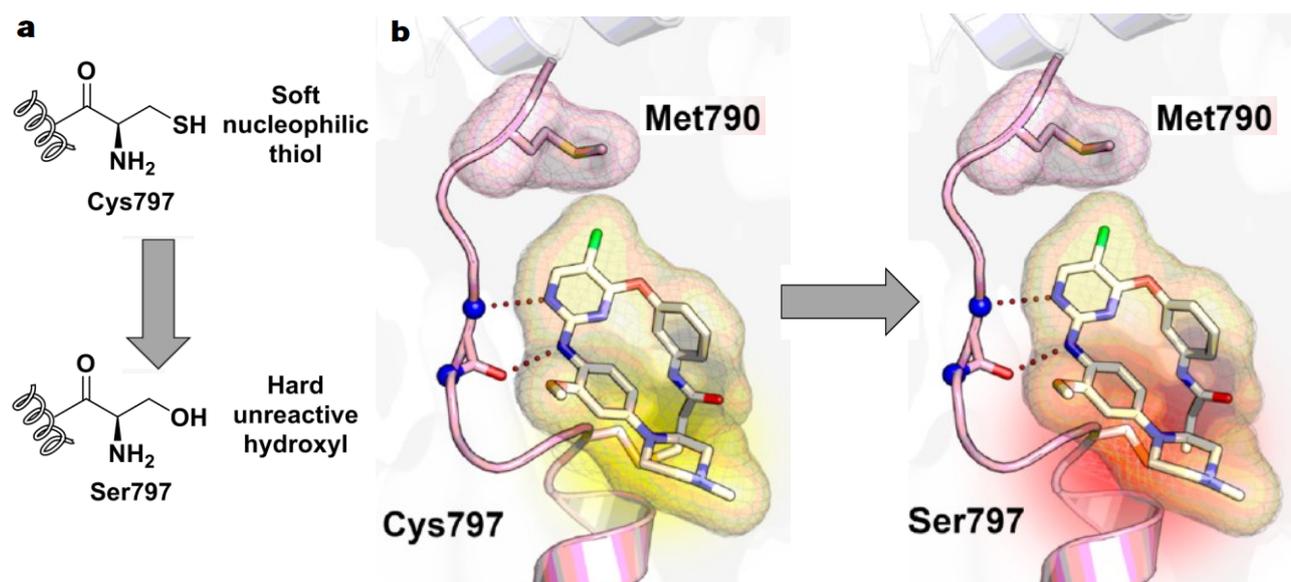


Figure 2.14: a) The EGFR C797S variant mediates resistance through point mutation of cysteine 797 to a serine residue and b) WZ4002 covalently bound to Cys797 highlighted in yellow and the less nucleophilic Ser797 unable to undergo covalent bond formation, highlighted in red (PDB: 3IKA). Figure reproduced from b) Engel et al.⁷⁷

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

Interestingly, the allelic/genomic context of resistance to 3rd generation inhibitors has been shown to influence the sensitivity of the receptor to the other EGFR TKIs available.⁸⁶ While C797S conferred resistance to all 3rd generation inhibitors, it was found to be sensitive to 1st and 2nd generation inhibitors such as gefitinib and afatinib. This may prove helpful in future therapeutic strategies for patients with EGFR mutations, paving the way for use of different sequential and combinatory treatment as the mechanism of acquired resistance varies based on the molecular therapeutic and context.

2.4.4 4th Generation Inhibitors and the Future of EGFR Inhibition

Since 2015, and the first evidence of 3rd generation inhibitor resistance mediated by EGFR-C797S, no major breakthroughs have been achieved to target the clinically relevant mutant variant. In the subsequent three years, various research groups have adopted novel approaches to attempt to overcome this mechanism of drug resistance. The focus has been on improving the reversible binding affinity of established aminopyrimidine-based inhibitors, investigating innovative hinge binding motifs and interactions with the mutant serine residue and allosteric inhibitors of the EGFR kinase domain. Compounds utilising these strategies have become known as 4th generation EGFR inhibitors.

The first of these strategies sought to improve the degree of reversible binding by exploiting a hydrophobic pocket adjacent to the gatekeeper residue, in combination with a hydrophobic clamp motif within the active site, as shown in **Figure 2.15a**, highlighted in green.⁸⁷ The lead compound, **SKLB(5)** (**Figure 2.15a**), is structurally similar to WZ4002, with the addition of a cyclopentyl moiety that interacts with the hydrophobic clamp and an aniline fragment which extends into the hydrophobic back pocket. Previously developed EGFR inhibitors have not made use of this hydrophobic pocket or clamp motif, providing favourable reversible interactions for the lead compound.

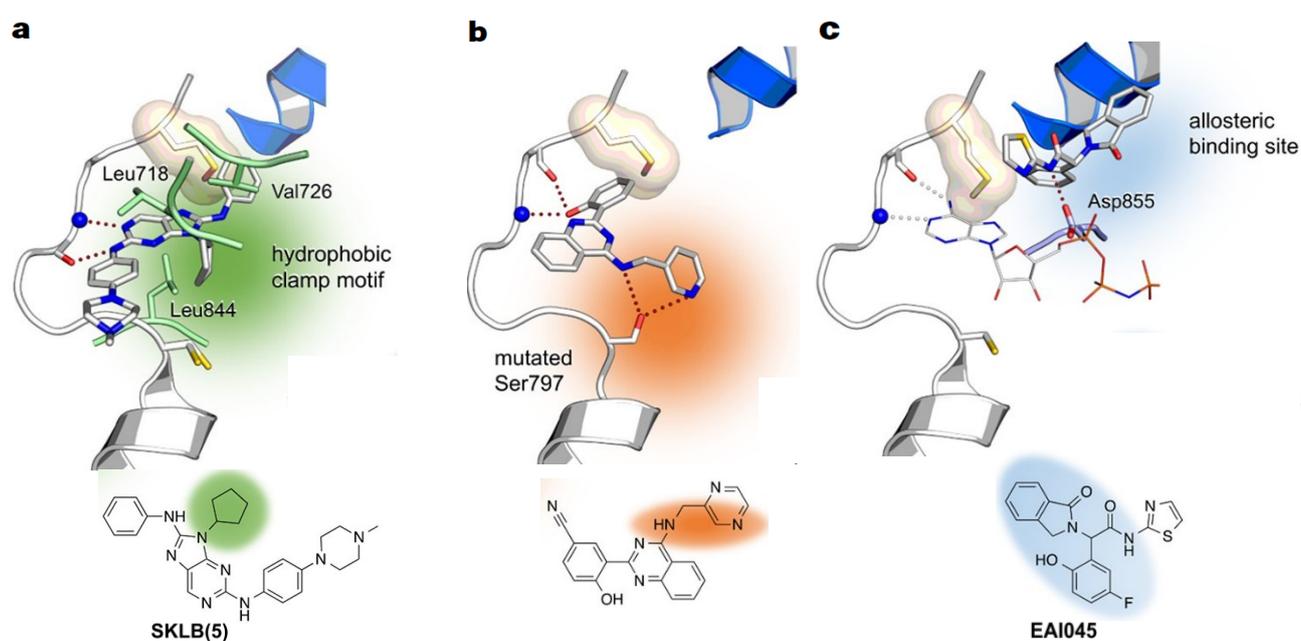


Figure 2.15: 4th Generation inhibitors and the various strategies undertaken to target the EGFR-C797S mutation. These include: a) Exploiting the hydrophobic pocket and clamp motif (green, PDB: 5GTZ), b) Reversible interaction with the mutated serine residue (orange, PDB: 5X2C), and c) Allosteric EGFR inhibitors (light blue, PDB: 5D41). Figure reproduced from Grabe et al.⁸⁸

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

Another approach utilised a virtual *de novo* design to identify a novel scaffold as a potent inhibitor for the del19/T790M/C797S triple mutant variant.⁸⁹ The most efficacious of these inhibitors, shown in **Figure 2.15b**, uses an unconventional hinge-binding motif, interacting with a phenolic hydroxyl moiety, as well as the aminomethylpyrazine forming two hydrogen bond interactions with Ser797 (**Figure 2.15b**, highlighted in orange).⁸⁹ This improved the binding affinity, while specifically exploiting the C797S mutant variant serine residue.

Lastly, the first of its kind allosteric EGFR inhibitor, **EAI045 (Figure 2.15c)**, was shown to overcome EGFR-C797S by targeting an allosteric pocket (**Figure 2.15c**, highlighted in blue) that emerges with the inactive conformation of the kinase.⁹⁰ The inhibitor showed low nanomolar activity in biochemical assays, whilst sparing the wild-type enzyme. However, this efficacy is only achievable on a cellular level or *in vivo* setting in combination with the antibody cetuximab.⁹⁰ Unfortunately, approaches based on EAI045 are limited to the L858R activating mutation, as these compounds show little to no efficacy against del19 activated EGFR, which is the predominant activating mutation found in osimertinib resistant patients.

While these methods of drug development have afforded numerous biochemically active 4th generation inhibitors, few have proven efficacious on a cellular level or *in vivo*. This reflects the current situation in the targeted inhibition of EGFR using TKIs. The emergence of the T790M mutation over a decade ago showed us that simply trying to improve on existing inhibitors, as was the case with the majority of 2nd generation inhibitors derived from gefitinib and erlotinib, does not necessarily counteract new mechanisms of acquired drug resistance. Novel chemical structures and binding modes, making use of the combined innovative approaches discussed above, will thus be required to unravel the molecular complexities of multimutated EGFR to develop 4th generation inhibitors for treatment of NSCLC patients. The primary goal of this body of research centres on the development, utilisation and evaluation of approaches to overcome the clinically relevant EGFR-C797S mutant variant. In the following chapter, the main aims and objectives of this work will be presented in summarised form, providing the reader with an overview of the strategies employed to overcome this resistance mechanism.

2.5 References

1. M. A. Lemmon and J. Schlessinger, *Cell*, 2010, **141**, 1117-1134.
2. R. Roskoski, *Pharmacological Research*, 2014, **79**, 34-74.
3. D. M. Thompson and G. N. Gill, *Cancer Surveys*, 1985, **4**, 767-788.
4. J. Marx, *Science*, 1986, **234**, 543-544.
5. S. Cohen, *Angewandte Chemie International Edition* 1987, **26**, 717-722.
6. K. M. Ferguson, *Annual Review of Biophysics*, 2008, **37**, 353-373.
7. D. Goodsell, *RCSB PDB-101: Molecule of the Month*, 2010.
8. X. Zhang, J. Gureasko, K. Shen, P. A. Cole and J. Kuriyan, *Cell*, 2006, **125**, 1137-1149.
9. N. Jura, N. F. Endres, K. Engel, S. Deindl, R. Das, M. H. Lamers, D. E. Wemmer, X. Zhang and J. Kuriyan, *Cell*, 2009, **137**, 1293-1307.
10. Y. Yarden and M. X. Sliwkowski, *Nature Reviews Molecular Cell Biology*, 2001, **2**, 127-137.
11. D. Bray and S. Lay, *Biophysical Journal*, 1994, **66**, 972-977.
12. R. Zandi, A. B. Larsen, P. Andersen, M.-T. Stockhausen and H. S. Poulsen, *Cellular Signalling*, 2007, **19**, 2013-2023.
13. R. S. Herbst, J. V. Heymach and S. M. Lippman, *New England Journal of Medicine*, 2008, **359**, 1367-1380.
14. R. I. Nicholson, J. M. W. Gee and M. E. Harper, *European Journal of Cancer*, 2001, **37**, 9-15.
15. E. Purba, E.-i. Saita and I. Maruyama, *Cells*, 2017, **6**, 1-19.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

16. J. H. Ludes-Meyers, M. A. Subler, C. V. Shivakumar, R. M. Munoz, P. Jiang, J. E. Bigger, D. R. Brown, S. P. Deb and S. Deb, *Molecular and Cellular Biology*, 1996, **16**, 6009-6019.
17. A. F. Gazdar and J. D. Minna, *Cancer Prevention Research*, 2008, **1**, 156-160.
18. D. B. Agus, R. W. Akita, W. D. Fox, G. D. Lewis, B. Higgins, P. I. Pisacane, J. A. Lofgren, C. Tindell, D. P. Evans, K. Maiese, H. I. Scher and M. X. Sliwkowski, *Cancer Cell*, 2002, **2**, 127-137.
19. W. Guo and F. G. Giancotti, *Nature Reviews Molecular Cell Biology*, 2004, **5**, 816-826.
20. K. G. Bache, T. Slagsvold and H. Stenmark, *The EMBO Journal*, 2004, **23**, 2707-2712.
21. K. Haglund, S. Sigismund, S. Polo, I. Szymkiewicz, P. P. Di Fiore and I. Dikic, *Nature Cell Biology*, 2003, **5**, 461-466.
22. C. T. Chu, K. D. Everiss, C. J. Wikstrand, S. K. Batra, H. J. Kung and D. D. Bigner, *Biochemical Journal*, 1997, **324**, 855-861.
23. M. V. Grandal, R. Zandi, M. W. Pedersen, B. M. Willumsen, B. van Deurs and H. S. Poulsen, *Carcinogenesis*, 2007, **28**, 1408-1417.
24. M. R. Brewer, S. H. Choi, D. Alvarado, K. Moravcevic, A. Pozzi, M. A. Lemmon and G. Carpenter, *Molecular Cell*, 2009, **34**, 641-651.
25. H. Shigematsu, L. Lin, T. Takahashi, M. Nomura, M. Suzuki, I. I. Wistuba, K. M. Fong, H. Lee, S. Toyooka, N. Shimizu, T. Fujisawa, Z. Feng, J. A. Roth, J. Herz, J. D. Minna and A. F. Gazdar, *Journal of the National Cancer Institute*, 2005, **97**, 339-346.
26. T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman and D. A. Haber, *New England Journal of Medicine*, 2004, **350**, 2129-2139.
27. S. Kobayashi, T. J. Boggon, T. Dayaram, P. A. Jänne, O. Kocher, M. Meyerson, B. E. Johnson, M. J. Eck, D. G. Tenen and B. Halmos, *New England Journal of Medicine*, 2005, **352**, 786-792.
28. E. H. Fischer and E. G. Krebs, *Journal of Biological Chemistry*, 1955, **216**, 121-132.
29. E. G. Krebs, *Angewandte Chemie International Edition*, 1993, **32**, 1122-1129.
30. E. H. Fischer, *Angewandte Chemie International Edition*, 1993, **32**, 1130-1137.
31. D. A. Walsh, J. P. Perkins and E. G. Krebs, *Journal of Biological Chemistry*, 1968, **243**, 3763-3765.
32. T. C. Linn, F. H. Pettit, F. Hucho and L. J. Reed, *Proceedings of the National Academy of Sciences of the United States of America*, 1969, **64**, 227-234.
33. P. Cohen, *Nature Cell Biology*, 2002, **4**, E127-E130.
34. S. Shoji, D. C. Parmelee, R. D. Wade, S. Kumar, L. H. Ericsson, K. A. Walsh, H. Neurath, G. L. Long, J. G. Demaille, E. H. Fischer and K. Titani, *Proceedings of the National Academy of Sciences of the United States of America*, 1981, **78**, 848-851.
35. M. S. Collett and R. L. Erikson, *Proceedings of the National Academy of Sciences of the United States of America*, 1978, **75**, 2021-2024.
36. T. Hunter and B. M. Sefton, *Proceedings of the National Academy of Sciences of the United States of America*, 1980, **77**, 1311-1315.
37. H. Ushiro and S. Cohen, *Journal of Biological Chemistry*, 1980, **255**, 8363-8365.
38. S. J. Leever and C. J. Marshall, *The EMBO journal*, 1992, **11**, 569-574.
39. N. Gómez and P. Cohen, *Nature*, 1991, **353**, 170-173.
40. G. Manning, D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam, *Science*, 2002, **298**, 1912-1934.
41. P. Lahiry, A. Torkamani, N. J. Schork and R. A. Hegele, *Nature Reviews Genetics*, 2010, **11**, 60-74.
42. J. A. Ubersax and J. E. Ferrell Jr, *Nature Reviews Molecular Cell Biology*, 2007, **8**, 530-541.
43. P. Warnault, A. Yasri, M. Coisy-Quivy, G. Cheve, C. Bories, B. Fauvel and R. Benhida, *Current Medicinal Chemistry*, 2013, **20**, 2043-2067.
44. A. K. Ghose, T. Herbertz, D. A. Pippin, J. M. Salvino and J. P. Mallamo, *Journal of Medicinal Chemistry*, 2008, **51**, 5149-5171.
45. R. Williams, A. Berndt, S. Miller, W.-C. Hon and X. Zhang, *Biochemical Society Transactions*, 2009, **37**, 615-626.
46. N. R. Brown, M. E. M. Noble, J. A. Endicott and L. N. Johnson, *Nature Cell Biology*, 1999, **1**, 438-443.
47. K. S. Bhullar, N. O. Lagarón, E. M. McGowan, I. Parmar, A. Jha, B. P. Hubbard and H. P. V. Rupasinghe, *Molecular Cancer*, 2018, **17**, 17-48.
48. S. Klaeger, S. Heinzlmeir, M. Wilhelm, H. Polzer, B. Vick, P.-A. Koenig, M. Reinecke, B. Ruprecht, S. Petzoldt, C. Meng, J. Zecha, K. Reiter, H. Qiao, D. Helm, H. Koch, M. Schoof, G. Canevari, E. Casale, S. R. Depaolini, A. Feuchtinger, Z. Wu, T. Schmidt, L. Rueckert, W. Becker, J. Huenges, A.-K. Garz, B.-O. Gohlke, D. P. Zolg, G. Kayser, T. Vooder, R. Preissner, H. Hahne, N. Tönisson, K. Kramer, K. Götze, F.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

- Bassermann, J. Schlegl, H.-C. Ehrlich, S. Aiche, A. Walch, P. A. Greif, S. Schneider, E. R. Felder, J. Ruland, G. Médard, I. Jeremias, K. Spiekermann and B. Kuster, *Science*, 2017, **358**, 1-16.
49. P. Wu, T. E. Nielsen and M. H. Clausen, *Trends in Pharmacological Sciences*, 2015, **36**, 422-439.
50. R. Roskoski, *Pharmacological Research*, 2016, **103**, 26-48.
51. L. Vandana and G. Indraneel, *Current Pharmaceutical Design*, 2012, **18**, 2936-2945.
52. Q. Liu, Y. Sabnis, Z. Zhao, T. Zhang, Sara J. Buhrlage, Lyn H. Jones and Nathanael S. Gray, *Chemistry & Biology*, 2013, **20**, 146-159.
53. Z. Zhao, Q. Liu, S. Bliven, L. Xie and P. E. Bourne, *Journal of Medicinal Chemistry*, 2017, **60**, 2879-2889.
54. J. M. Bradshaw, J. M. McFarland, V. O. Paavilainen, A. Bisconte, D. Tam, V. T. Phan, S. Romanov, D. Finkle, J. Shu, V. Patel, T. Ton, X. Li, D. G. Loughhead, P. A. Nunn, D. E. Karr, M. E. Gerritsen, J. O. Funk, T. D. Owens, E. Verner, K. A. Brameld, R. J. Hill, D. M. Goldstein and J. Taunton, *Nature Chemical Biology*, 2015, **11**, 525-531.
55. E. Leproult, S. Barluenga, D. Moras, J.-M. Wurtz and N. Winssinger, *Journal of Medicinal Chemistry*, 2011, **54**, 1347-1355.
56. R. S. Herbst, M. Fukuoka and J. Baselga, *Nature Reviews Cancer*, 2004, **4**, 979-987.
57. P. Seshacharyulu, M. P. Ponnusamy, D. Haridas, M. Jain, A. K. Ganti and S. K. Batra, *Expert Opinion on Therapeutic Targets*, 2012, **16**, 15-31.
58. M. H. Cohen, J. R. Johnson, Y.-F. Chen, R. Sridhara and R. Pazdur, *The Oncologist*, 2005, **10**, 461-466.
59. P. J. Medina and S. Goodin, *Clinical Therapeutics*, 2008, **30**, 1426-1447.
60. A. Mullard, *Nature Reviews Drug Discovery*, 2012, **11**, 91-94.
61. J. G. Paez, P. A. Jänne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M. J. Eck, W. R. Sellers, B. E. Johnson and M. Meyerson, *Science*, 2004, **304**, 1497-1500.
62. D. M. Jackman, B. Y. Yeap, L. V. Sequist, N. Lindeman, A. J. Holmes, V. A. Joshi, D. W. Bell, M. S. Huberman, B. Halmos, M. S. Rabin, D. A. Haber, T. J. Lynch, M. Meyerson, B. E. Johnson and P. A. Jänne, *Clinical Cancer Research*, 2006, **12**, 3908-3914.
63. A. Russo, T. Franchina, G. R. Ricciardi, A. Picone, G. Ferraro, Z. Mariangela, G. Toscano, A. Giordano and V. Adamo, *Oncotarget*, 2015, **6**, 26814-26825.
64. C.-H. Yun, T. J. Boggon, Y. Li, M. S. Woo, H. Greulich, M. Meyerson and M. J. Eck, *Cancer Cell*, 2007, **11**, 217-227.
65. J.-C. Soria, T. S. Mok, F. Cappuzzo and P. A. Jänne, *Cancer Treatment Reviews*, 2012, **38**, 416-430.
66. W. Pao, V. A. Miller, K. A. Politi, G. J. Riely, R. Somwar, M. F. Zakowski, M. G. Kris and H. Varmus, *PLOS Medicine*, 2005, **2**, 225-235.
67. D. Westover, J. Zugazagoitia, B. C. Cho, C. M. Lovly and L. Paz-Ares, *Annals of Oncology*, 2018, **29**, i10-i19.
68. C.-H. Yun, K. E. Mengwasser, A. V. Toms, M. S. Woo, H. Greulich, K.-K. Wong, M. Meyerson and M. J. Eck, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 2070-2075.
69. J. Chmielecki, J. Foo, G. R. Oxnard, K. Hutchinson, K. Ohashi, R. Somwar, L. Wang, K. R. Amato, M. Arcila, M. L. Sos, N. D. Socci, A. Viale, E. de Stanchina, M. S. Ginsberg, R. K. Thomas, M. G. Kris, A. Inoue, M. Ladanyi, V. A. Miller, F. Michor and W. Pao, *Science Translational Medicine*, 2011, **3**, 1-10.
70. K. Park, E.-H. Tan, K. O'Byrne, L. Zhang, M. Boyer, T. Mok, V. Hirsh, J. C.-H. Yang, K. H. Lee, S. Lu, Y. Shi, S.-W. Kim, J. Laskin, D.-W. Kim, C. D. Arvis, K. Kölblbeck, S. A. Laurie, C.-M. Tsai, M. Shahidi, M. Kim, D. Massey, V. Zazulina and L. Paz-Ares, *The Lancet Oncology*, 2016, **17**, 577-589.
71. L. V. Sequist, J. C.-H. Yang, N. Yamamoto, K. O'Byrne, V. Hirsh, T. Mok, S. L. Geater, S. Orlov, C.-M. Tsai, M. Boyer, W.-C. Su, J. Bennouna, T. Kato, V. Gorbunova, K. H. Lee, R. Shah, D. Massey, V. Zazulina, M. Shahidi and M. Schuler, *Journal of Clinical Oncology*, 2013, **31**, 3327-3334.
72. R. T. Dunto and G. M. Keating, *Drugs*, 2013, **73**, 1503-1515.
73. D. Li, L. Ambrogio, T. Shimamura, S. Kubo, M. Takahashi, L. R. Chirieac, R. F. Padera, G. I. Shapiro, A. Baum, F. Himmelsbach, W. J. Rettig, M. Meyerson, F. Solca, H. Greulich and K. K. Wong, *Oncogene*, 2008, **27**, 4702-4711.
74. V. A. Miller, V. Hirsh, J. Cadranel, Y.-M. Chen, K. Park, S.-W. Kim, C. Zhou, W.-C. Su, M. Wang, Y. Sun, D. S. Heo, L. Crino, E.-H. Tan, T.-Y. Chao, M. Shahidi, X. J. Cong, R. M. Lorence and J. C.-H. Yang, *The Lancet Oncology*, 2012, **13**, 528-538.
75. L. V. Sequist, B. Besse, T. J. Lynch, V. A. Miller, K. K. Wong, B. Gitlitz, K. Eaton, C. Zacharchuk, A. Freyman, C. Powell, R. Ananthkrishnan, S. Quinn and J.-C. Soria, *Journal of Clinical Oncology*, 2010, **28**, 3076-3083.

Chapter 2 – Epidermal Growth Factor Receptor and its Targeted Inhibition

76. K. L. Reckamp, G. Giaccone, D. R. Camidge, S. M. Gadgeel, F. R. Khuri, J. A. Engelman, M. Koczywas, A. Rajan, A. K. Campbell, D. Gernhardt, A. Ruiz-Garcia, S. Letrent, J. Liang, I. Taylor, J. P. O'Connell and P. A. Jänne, *Cancer*, 2014, **120**, 1145-1154.
77. J. Engel, J. Lategahn and D. Rauh, *ACS Medicinal Chemistry Letters*, 2016, **7**, 2-5.
78. W. Zhou, D. Ercan, L. Chen, C.-H. Yun, D. Li, M. Capelletti, A. B. Cortot, L. Chirieac, R. E. Jacob, R. Padera, J. R. Engen, K.-K. Wong, M. J. Eck, N. S. Gray and P. A. Jänne, *Nature*, 2009, **462**, 1070-1074.
79. L. V. Sequist, J.-C. Soria, J. W. Goldman, H. A. Wakelee, S. M. Gadgeel, A. Varga, V. Papadimitrakopoulou, B. J. Solomon, G. R. Oxnard, R. Dziadziuszko, D. L. Aisner, R. C. Doebele, C. Galasso, E. B. Garon, R. S. Heist, J. Logan, J. W. Neal, M. A. Mendenhall, S. Nichols, Z. Piotrowska, A. J. Wozniak, M. Raponi, C. A. Karlovich, S. Jaw-Tsai, J. Isaacson, D. Despaigne, S. L. Matheny, L. Rolfe, A. R. Allen and D. R. Camidge, *New England Journal of Medicine*, 2015, **372**, 1700-1709.
80. N. Van Der Steen, C. Caparello, C. Rolfo, P. Pauwels, G. J. Peters and E. Giovannetti, *OncoTargets and Therapy*, 2016, **9**, 6065-6074.
81. K.-O. Lee, M. Y. Cha, M. Kim, J. Y. Song, J.-H. Lee, Y. H. Kim, Y.-M. Lee, K. H. Suh and J. Son, *Cancer Research*, 2014, **74**, LB-100.
82. A. Yver, *Annals of Oncology*, 2016, **27**, 1165-1170.
83. D. A. E. Cross, S. E. Ashton, S. Ghiorghiu, C. Eberlein, C. A. Nebhan, P. J. Spitzler, J. P. Orme, M. R. V. Finlay, R. A. Ward, M. J. Mellor, G. Hughes, A. Rahi, V. N. Jacobs, M. R. Brewer, E. Ichihara, J. Sun, H. Jin, P. Ballard, K. Al-Kadhimi, R. Rowlinson, T. Klinowska, G. H. P. Richmond, M. Cantarini, D.-W. Kim, M. R. Ranson and W. Pao, *Cancer Discovery*, 2014, **4**, 1046-1061.
84. K. S. Thress, C. P. Paweletz, E. Felip, B. C. Cho, D. Stetson, B. Dougherty, Z. Lai, A. Markovets, A. Vivancos, Y. Kuang, D. Ercan, S. E. Matthews, M. Cantarini, J. C. Barrett, P. A. Jänne and G. R. Oxnard, *Nature Medicine*, 2015, **21**, 1-5.
85. J. A. Woyach, R. R. Furman, T.-M. Liu, H. G. Ozer, M. Zapatka, A. S. Ruppert, L. Xue, D. H.-H. Li, S. M. Steggerda, M. Versele, S. S. Dave, J. Zhang, A. S. Yilmaz, S. M. Jaglowski, K. A. Blum, A. Lozanski, G. Lozanski, D. F. James, J. C. Barrientos, P. Lichter, S. Stilgenbauer, J. J. Buggy, B. Y. Chang, A. J. Johnson and J. C. Byrd, *New England Journal of Medicine*, 2014, **370**, 2286-2294.
86. M. J. Niederst, H. Hu, H. E. Mulvey, E. L. Lockerman, A. R. Garcia, Z. Piotrowska, L. V. Sequist and J. A. Engelman, *Clinical Cancer Research*, 2015, **21**, 3924-3933.
87. S.-J. Zhu, P. Zhao, J. Yang, R. Ma, X.-E. Yan, S.-Y. Yang, J.-W. Yang and C.-H. Yun, *Oncotarget*, 2018, **9**, 13652-13665.
88. T. Grabe, J. Lategahn and D. Rauh, *ACS Medicinal Chemistry Letters*, 2018, **9**, 779-782.
89. H. Park, H.-Y. Jung, S. Mah and S. Hong, *Angewandte Chemie International Edition*, 2017, **56**, 7634-7638.
90. Y. Jia, C.-H. Yun, E. Park, D. Ercan, M. Manuia, J. Juarez, C. Xu, K. Rhee, T. Chen, H. Zhang, S. Palakurthi, J. Jang, G. Lelais, M. DiDonato, B. Bursulaya, P.-Y. Michellys, R. Epple, T. H. Marsilje, M. McNeill, W. Lu, J. Harris, S. Bender, K.-K. Wong, P. A. Jänne and M. J. Eck, *Nature*, 2016, **534**, 129-132.

Chapter 3

Overview of Aims, Objectives and General Practices

3.1 Introduction

The intention of this chapter is to provide the reader with an overview of the various research themes, aims and objectives of each chapter within this thesis. This will be achieved by briefly introducing the main research goals of each chapter, providing the premise for the undertaking of these ventures and summarizing the synthetic strategies employed to accomplish these endeavours. Additionally, general practices regarding the synthesis and biochemical and cellular evaluation of all compounds will be comprehensively discussed. All figures displayed within this chapter were appropriated from the respective chapter under summation.

3.2 Chapter 4: Lead Optimization of Pyrrolopyrimidine-derived EGFR Inhibitors

In the ensuing Chapter 4, we will attempt to develop potent irreversible and reversible inhibitors of the cysteine bearing L858R/T790M double mutant and serine bearing L858R/T790M/C797S triple mutant respectively. To do this, we will utilise and optimise the pyrrolopyrimidine-based lead compound **1**, illustrated in **Figure 3.1** on the following page. This compound was previously identified and synthesized by Mr. Philip Klövekorn, a student under the supervision of our research collaborator, Prof. Daniel Rauh. As such, all of the synthetic work pertaining to this chapter was completed in the laboratories of Prof. Rauh during a funded research visit to Dortmund, Germany.

Following a discussion on the development and optimization towards lead compound **1**, we will introduce the dual EGFR/HER2 inhibitors which provided the inspiration for this project. These drug candidates contain glycol chain moieties of various lengths and substituents which were found to instate favourable reversible interactions with the glycine-rich loop within the ATP binding site.¹ Wishing to emulate the enhanced reversible binding affinity afforded by these groups, and aided by the use of molecular modelling, we envisaged the incorporation of glycol chains at strategic positions on lead compound **1**. This included derivatisation at the *N*-position (**Figure 3.1a**, R₁) and *ortho*-position (**Figure 3.1b**, R₂) of the acrylamide Michael acceptor with the selected glycol fragments (**Figure 3.1**, blue). Thus, we would be able to examine the effects of substitution at these positions with respect to binding affinity and formation of reversible interactions. In conjunction with this investigation, we would synthesise analogues of these compounds with, and without, the electrophilic acrylamide Michael acceptor (**Figure 3.1**, red). In doing so, we could determine their potential as irreversible inhibitors of the T790M double mutant and as reversible inhibitors of the C797S triple mutant respectively. Additionally, we would consider the necessity of the phenyl methyl-piperazine solubilising group (**Figure 3.1**, green) and the effect of its presence or absence on the cellular permeability and efficacy of these compounds. The derivatisation and interchange of these specific structural elements within the lead compound scaffold **1** will constitute the compounds displayed within the eight Series A – H, illustrated in **Figure 3.1**. Following the successful synthesis of this combinatorial library, these compounds will undergo biochemical and cellular evaluation against both the wild-type and mutant variants of the EGFR enzyme to determine these effects.

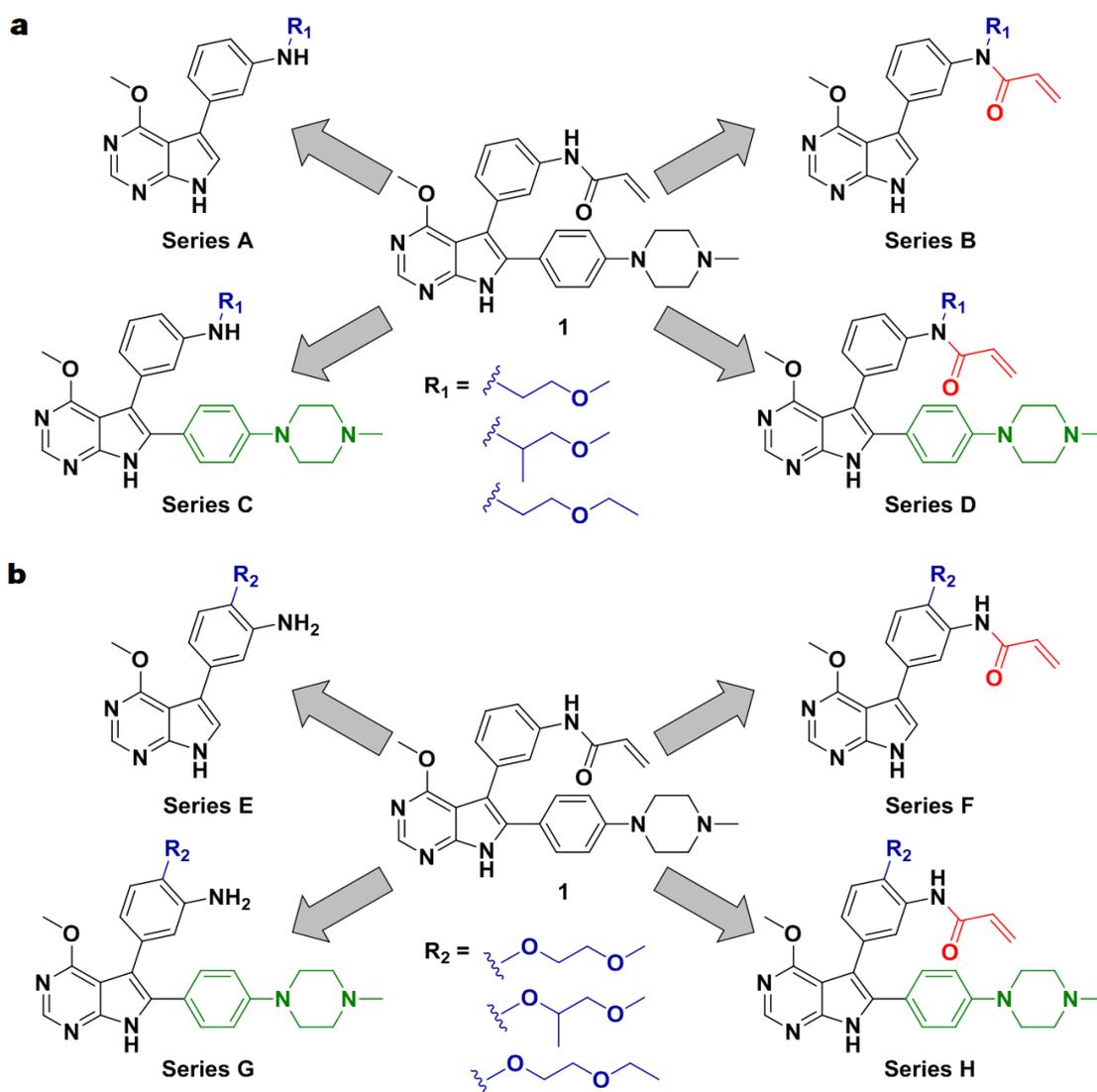


Figure 3.1: Structure of lead compound **1** and the combinatorial synthetic library containing a) Series A – D, with the glycol chain on the *N*-position of the acrylamide and b) Series E – H, with the glycol chain in an *ortho*-position to the acrylamide.

3.3 Chapter 5: 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Chapter 5 represents a departure from the traditional use of reversible and irreversible acrylamide electrophile-containing inhibitors which target cysteine residues within the active site. Instead, in the hopes of effectively combating resistance conferred by the C797S mutation, we will attempt to restore covalent inhibition in this mutant variant by targeting the EGFR catalytic lysine residue. The EGFR catalytic Lys745 residue plays a significant regulatory role in activation of the kinase, a prerequisite state and conformation which facilitates successful binding and phosphorylation of ATP and downstream signalling.² However, mutation of this amino acid has been documented to create an inactive or “kinase-dead” form of the enzyme.³ Therefore, we found the targeted inhibition of this specific residue an attractive prospect in a therapeutic context, with the potential to switch off constitutively activated oncogenic signalling.

Chapter 3 – Overview of Aims, Objectives and General Practices

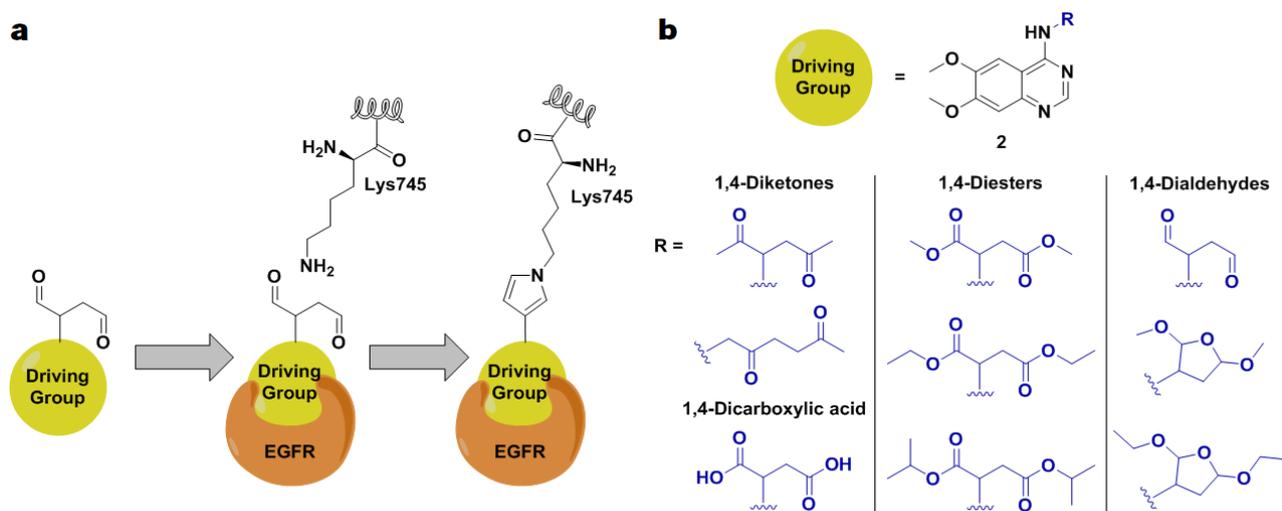


Figure 3.2: a) Use of a 1,4-dicarbonyl electrophile to potentially undergo Paal-Knorr pyrrole synthesis with the EGFR catalytic Lys745 residue and b) Illustration of the selected heterocyclic driving group **2** and various 1,4-dicarbonyl electrophiles constituting the prospective synthetic library.

Amongst several successful instances of lysine-targeting electrophiles that will be examined, we drew inspiration from our collaborator from Texas State University, Prof. Alexander Kornienko. His group was able to synthesise and investigate a series of analogues of the 1,4-dialdehyde-containing natural product polygodial, which displayed potent antiproliferative activity.⁴ The group hypothesised, and were able to successfully prove, that the dialdehyde moiety of the natural product underwent a modified Paal-Knorr condensation reaction with a lysine residue on the target protein.⁴ Seeking to reproduce this novel mechanism of covalent modification, we envisioned the incorporation of a 1,4-dicarbonyl electrophile on a suitable heterocyclic driving group to potentially undergo Paal-Knorr pyrrole synthesis with the EGFR catalytic Lys745 residue, as illustrated in **Figure 3.2a**.

To investigate the feasibility and possibility of irreversibly inhibiting the catalytic lysine residue, we envisage the synthesis of a library of 1,4-dicarbonyl-containing inhibitors. Through structural analysis of scaffolds bound to the EGFR kinase domain and use of molecular modelling, the gefitinib-derived, quinazoline core **2**, shown at the top right of **Figure 3.2b**, will be selected as our heterocyclic driving group. We will employ a variety of 1,4-dicarbonyl electrophiles (**Figure 3.2b**, blue), to compare the effects of respective electrophile reactivity, steric bulk/leaving groups, lipophilicity and stability of the formed covalent adducts on the promotion and rate of the Paal-Knorr condensation reaction. Following the successful synthesis of this library of compounds, these comparisons and the binding mode of the inhibitors will be explored through the results obtained from the biochemical and cellular assessment and covalent mass spectrometry experiments respectively.

3.4 Chapter 6: Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Advancing on the premise of the previous chapter, we will pursue irreversible inhibition of the catalytic Lys745 residue further throughout Chapter 6. However, at the onset of this research endeavour, Taunton and co-workers reported the first crystal structure of EGFR in which the catalytic Lys745 residue was covalently modified.⁵ As the group made use of a broad-spectrum chemical probe with excellent reversible binding affinity and a sulfonyl fluoride electrophile to undergo covalent modification with the catalytic Lys745 residue, we were compelled to re-evaluate the strategy implemented in the previous chapter. Therefore, we sought to identify an optimised and selective EGFR heterocyclic driving group to combine with reactive lysine-targeting electrophiles such as the sulfonyl fluoride employed in the publication by Zhao et al.⁵

Experimental evidence within the literature, coupled with the incredible success and efficacy of osimertinib as a 3rd generation EGFR inhibitor, led to its selection as the driving group scaffold for our synthetic library.⁶ This scaffold **3**, pictured in **Figure 3.3** below, utilises the potent 2-aminopyrimidine group (green) as hinge binding motif, which has been documented to overcome resistance in both ALK and EGFR.⁷ For incorporation of the lysine-targeting electrophiles, the optimal position was identified as the *N*-position (**Figure 3.3**, R₁, blue) of the indole heterocycle through use of molecular modelling. This will allow for a proximal location and favourable orientation towards the Lys745 residue for the proposed electrophiles. These functionalities (**Figure 3.3**, blue), are chosen for their ability to form favourable reversible interactions or covalently modify lysine residues, which we hope will improve the efficacy against and selectivity for enzymes bearing the C797S mutation.⁸

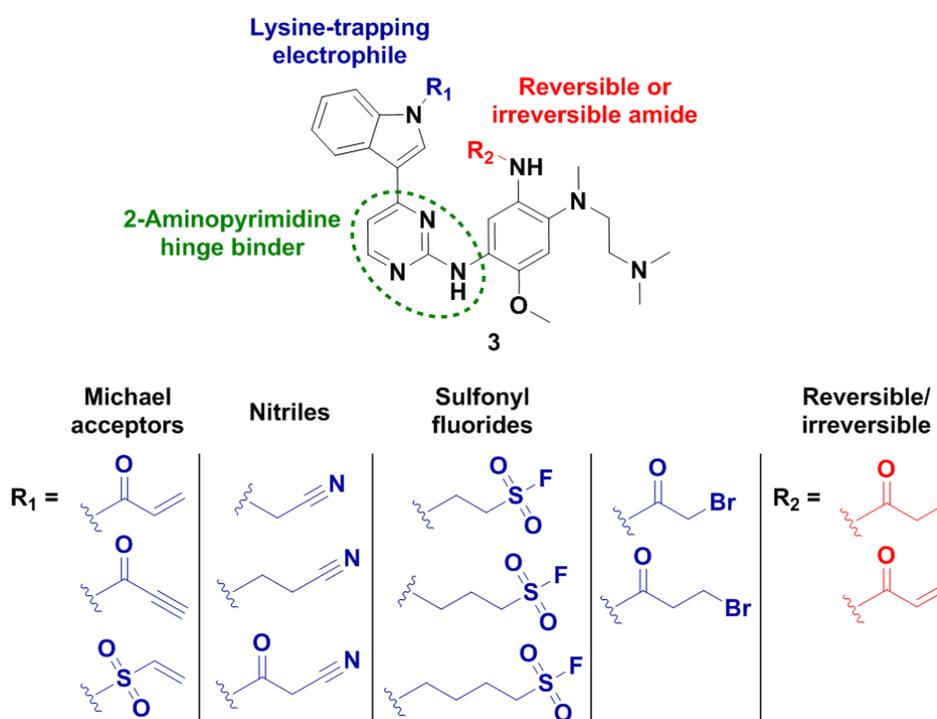


Figure 3.3: Target synthetic library for osimertinib-derived catalytic lysine-targeting inhibitors.

Chapter 3 – Overview of Aims, Objectives and General Practices

Lastly, we wish to synthesise the reversible propionamide and irreversible acrylamide analogues (**Figure 3.3**, R₂, red) of our osimertinib-derived inhibitors. This could provide valuable insight and a comparative platform in the case of covalent bond formation at the catalytic lysine residue. Once again, the potency and binding mode of these synthesised inhibitors will be investigated according to the outcome of the biochemical and cellular evaluation and covalent mass spectrometry experiments respectively.

3.5 Chapter 7: Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

While the research efforts of the previous two chapters were centred on the covalent modification and targeted inhibition of the EGFR catalytic Lys745 residue, in our final chapter we seek to exploit and irreversibly inhibit the main mechanism of resistance to 3rd generation EGFR inhibitors - the mutated Ser797 residue.⁹ In doing so, we would reinstate irreversible inhibition and produce selective inhibitors of the C797S mutation-bearing enzyme.

As this serine residue contains a harder, less nucleophilic hydroxyl group, our strategy will include the use of suitable electrophiles known to interact and irreversibly inhibit serine-containing enzymes. Building on the synthetic achievements of Chapter 6, and exploiting the favourable attributes of this driving group, these electrophiles will be installed onto the previously implemented osimertinib-derived scaffold **4**, displayed in **Figure 3.4**. Derivatisation will take place at the aniline fragment (**Figure 3.4**, R₁, purple) which conventionally houses the acrylamide Michael acceptor, thereby placing the warhead in a prime position to interact and potentially covalently modify the mutant Ser797 residue. We will employ numerous electrophiles (**Figure 3.4**, purple) with varying reactivities to give the highest probability for irreversible inhibition of the EGFR-C797S mutant variant.

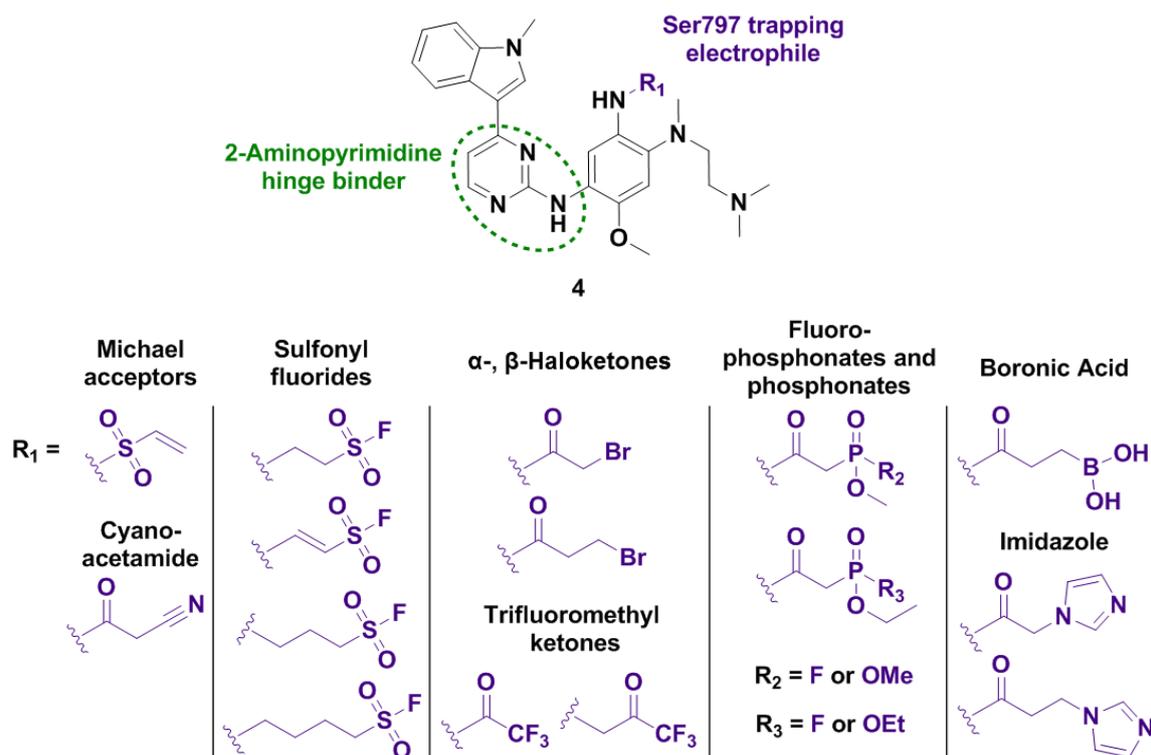


Figure 3.4: Target synthetic library for osimertinib-derived Ser797 targeting inhibitors.

Chapter 3 – Overview of Aims, Objectives and General Practices

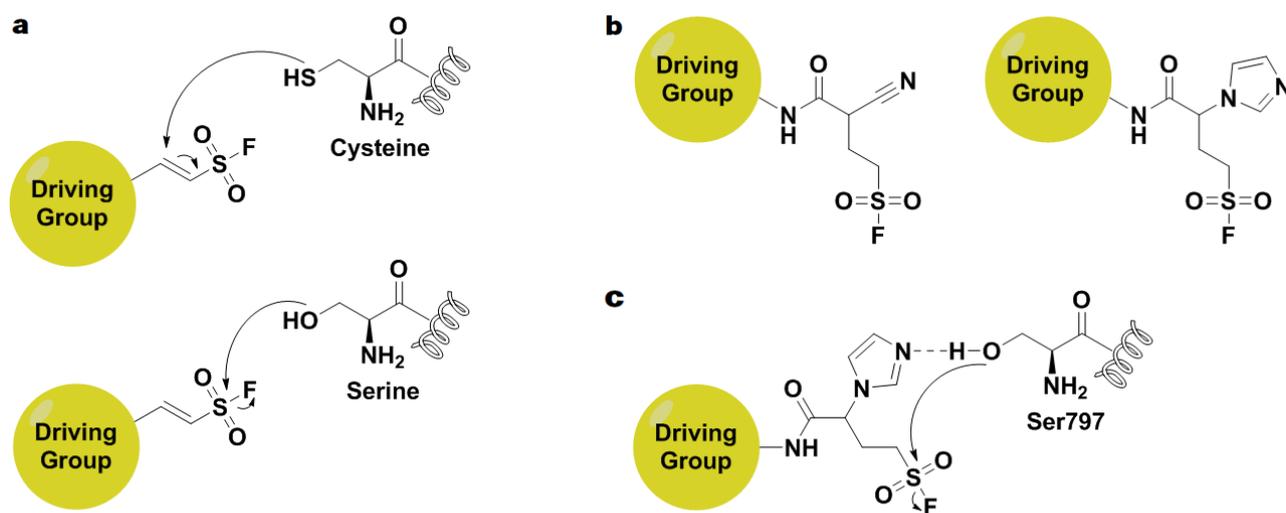


Figure 3.5: a) Dual warhead able to potentially covalently modify both cysteine and serine, b) Envisaged dual warheads and c) Mechanism of activation based on the catalytic triad of serine proteases.

Additionally, we wish to synthesise the dual-warhead-bearing compounds, illustrated above in **Figure 3.5**, which may show enhanced selectivity and efficiency in targeting the mutant Ser797 residue. These dual-warheads will include the vinyl sulfonyl fluoride (**Figure 3.5a**), able to potentially covalently modify both cysteine and serine residues through the Michael acceptor and adjacent sulfonyl centre respectively, and the cyano- and imidazole-containing inhibitors (**Figure 3.5b**). The latter inhibitors, which will be based on the catalytic triad of serine proteases, will be able to potentially activate the Ser797 residue through hydrogen bond interactions with the imidazole (or cyano) moiety. This can then be followed by potential irreversible inhibition through nucleophilic attack of the sulfonyl fluoride centre (**Figure 3.5c**).

As with all our synthesised libraries, these compounds will be assessed in a biochemical and cellular setting and undergo covalent mass spectrometry experiments to reveal the binding mode of the inhibitors. Special consideration will be given to the results of enzymes harbouring the C797S mutation, as we wish to selectively and effectively inhibit this mutant variant.

3.6 Chapter 8: Project Summary and Concluding Remarks

In the concluding chapter of this thesis we will provide commentary on the project as a collective and the future of EGFR as a therapeutic target. This will include a discussion of the present situation and outlook of EGFR inhibition and its relevance in the clinical treatment of cancer patients, a summary of the achievements and outcomes of this research project and the prospective application of this work.

3.7 General Practices

3.7.1 Synthesis

All compounds furnished and described within this thesis were synthesised by Luke Hodson.

3.7.1.1 Solvents and Reagents

Chemicals, starting reagents and solvents used in these experiments were purchased from VWR, Alfa Aesar, Acros Organics, Sigma Aldrich or Merck. Tetrahydrofuran, diethyl ether and toluene were distilled under nitrogen from sodium wire/sand using benzophenone as indicator. Dichloromethane, dimethylformamide and acetonitrile were distilled under nitrogen from calcium hydride. Methanol, ethanol and isopropanol were distilled under nitrogen from magnesium turnings and catalytic iodine. Alternatively, solvents were dried for at least two days in a sealed Schlenk flask, under argon, using 3 Å, 4 Å or 5 Å molecular sieves. Triethylamine and *N,N*-diisopropylethylamine were distilled under nitrogen from potassium hydroxide. Ethyl acetate, petroleum ether and dichloromethane used for flash column chromatography was distilled under open air conditions in bulk batches. All other reagents that required purification were purified according to standard procedures.¹⁰

3.7.1.2 Chromatography and Purification

Thin layer chromatography (TLC) was performed on Macherey Nagel aluminium TLC-plates, pre-coated with 0.20 mm silica gel and fluorescent indicator UV₂₅₄. Visualization was performed with UV light ($\lambda = 254$ nm), iodine on silica or by spraying with KMnO₄, *p*-anisaldehyde, ceric ammonium molybdate (CAM), ninhydrin (NIN), vanillin or bromocresol green stains followed by heating. Universal pH test paper strips were used to determine pH. Flash column chromatography was performed using Merck silica gel 60 (particle size 0.040-0.063 mm) or neutral alumina. Mobile phase and *R_f* value solvent ratios are reported as *v/v* ratios. Automated flash column chromatography was carried out on a Biotage Isolera using Reverleris silica flash cartridges (25 – 120 g) monitored by UV at $\lambda = 254$ nm and 366 nm. A rotary evaporator was used to remove solvents *in vacuo*. High vacuum (~ 0.8 mm Hg) was used to dry products. All final compounds were purified to >95 % purity as determined by high-performance liquid chromatography (HPLC). Purity was measured using a Waters Acquity UPLC system fitted with a photodiode array detector (system: Waters BEH C18 column (1.7 μ m, 100 mm \times 2.1 mm) by the Central Analytical Facility of Stellenbosch University. Alternatively, purity was measured using an Agilent 1200 series HPLC system with UV detection at $\lambda = 210$ nm (system: Agilent Eclipse XDB-C18 4.6 mm \times 150 mm, 5 μ M, 10 – 100 % CH₃CN in H₂O, with 0.1 % TFA, for 15 min at 1.00 mL/min) at the Technische Universität Dortmund.

Chapter 3 – Overview of Aims, Objectives and General Practices

3.7.1.3 Spectroscopic Characterization and Physical Properties

All infrared spectra were recorded on a Thermo Nicolet FT-IR, using an Attenuated Total Reflectance (ATR) attachment. OMNIC 7.0 software was used to analyse spectra. NMR spectra were recorded on a 300 MHz Varian VNMRS (75 MHz for ^{13}C NMR spectra, 96 MHz for ^{11}B NMR spectra, 282 MHz for ^{19}F NMR spectra and 121 MHz for ^{31}P NMR spectra), 400 MHz Varian Unity Inova (101 MHz for ^{13}C NMR spectra, 162 MHz for ^{31}P NMR spectra) or 600 MHz Varian Unity Inova (151 MHz for ^{13}C NMR spectra) at the Central Analytical Facilities (CAF) of Stellenbosch University. Alternatively, NMR spectra were recorded on a 400 MHz Bruker Avance-III HD (101 MHz for ^{13}C NMR spectra), 500 MHz Bruker Avance DRX (126 MHz for ^{13}C NMR spectra), 600 MHz Bruker Avance-III HD (151 MHz for ^{13}C NMR spectra) or 700 MHz Bruker Avance DRX (176 MHz for ^{13}C NMR spectra) at the Technische Universität Dortmund. All NMR spectra were obtained at 25 °C, chemical shifts (δ) are reported in parts per million (ppm), multiplicities are indicated as s (singlet), d (doublet), dd (doublet of doublets), t (triplet), q (quartet), m (multiplet) and br s (broad singlet) and coupling constants (J) are expressed in Hertz (Hz). ^1H and ^{13}C spectra are referenced to the residual solvent signal DMSO- d_6 (2.50 or 39.52 ppm), CDCl_3 (7.26 or 77.16 ppm) or CD_3OD (1.94 or 49.00 ppm) respectively. ^{11}B , ^{19}F and ^{31}P NMR spectra are referenced to 15% $\text{BF}_3\cdot\text{OEt}_2$ in CDCl_3 , CFCl_3 and 85% H_3PO_4 in H_2O respectively, which served as the internal standard for these experiments. NMR spectra were processed using MestReNova version 6.0.2-5475. High-resolution mass spectrometry was performed using a Waters SYNAPT G2 QTOF spectrometer, in ESI positive mode, by the CAF at Stellenbosch University. Alternatively, high-resolution mass spectrometry was performed using a Thermo LTQ Orbitrap in ESI positive mode, at the Technische Universität Dortmund. Melting points were obtained using a Gallenkamp melting point apparatus or a Kofler microscope melting point machine.

3.7.1.4 Glassware, Inert Conditions and Temperature Control

Glassware was oven dried and thereafter was placed under vacuum of ~ 0.8 mm Hg and cyclically flushed with nitrogen/argon and evacuated until it had reached room temperature. Standard Schlenk techniques were employed when necessary. All reactions were performed under a positive pressure of 2.8 kPa of 5.0 grade nitrogen or argon (Air Products). Low temperature reactions were performed in a dewar containing ice in acetone (~ -10 °C), dry ice in acetonitrile (-40 °C) or dry ice in acetone (-78 °C). Reactions requiring precise, extended low temperature control, were performed in a dewar regulated with a Thermo Scientific Haake EK90 Immersion Cooler.

3.7.2 Biochemical and Cellular Evaluation

All biochemical and cellular evaluations, as well as the preparative procedures described below, were undertaken by our collaborators in the research group of Prof. Daniel Rauh at the Technische Universität Dortmund.

3.7.2.1 Protein Expression and Purification

EGFR-T790M and EGFR-L858R/T790M/C797S were expressed in *Spodoptera frugiperda* (Sf9) cells using the BacMagic system and purified as follows. After three days of expression (27 °C, 110 rpm) the cells were harvested (300 g, 20 min), resuspended in buffer A (50 mM Tris, 500 mM NaCl, 5% glycerol, 1 mM DTT, pH 7.5) and homogenized by French press. The lysate was cleared by centrifugation at 40.000 g for 1 h at 8 °C and loaded on a pre-packed GSTrap 4B Column (GE Healthcare). The elution was done with buffer B (50 mM Tris, 500 mM NaCl, 5% glycerol, 1 mM DTT, pH 7.5, 10 mM glutathione). Fractions containing the assay construct EGFR-L858R/T790M/C797S were then concentrated and applied to a HiLoad 16/600 superdex 75 pg column (GE Healthcare) in buffer C (25 mM TRIS, 250 mM NaCl, 10% Glycerol, pH 8). The purified protein was concentrated to 5 mg/mL and stored at –80 °C until further use. Protein identity was confirmed by ESI-MS analysis.

3.7.2.2 Construct Design of EGFR-L858R/T790M/C797S

For the EGFR-L858R/T790M/C797S assay construct the DNA encoding residues comprising the juxtamembrane segment, the kinase domain and the C-terminal tail of human EGFR (UniProt entry P00533, residues 695-1210) were synthesized by GeneArt (Life Technologies). The construct was cloned into pLEX/Bac5 expression vector (MerckMillipore) including a GST-tag, using KpnI and Bsu36I restriction sites. Point mutations L858R, T790M and C797S were introduced by site-directed mutagenesis (QuikChange, Stratagene/Agilent Technologies). Transfection, virus generation and amplification were carried out in a Sf9 cell line following the BacMagic protocol.

3.7.2.3 Activity-based Assay for IC₅₀ Determination

IC₅₀ determinations for EGFR and its mutants (Carna Biosciences, lot13CBS-0005K for EGFR-wt; Carna, lot13CBS-0537B for EGFR-L858R and Carna, lot12CBS-0765B for EGFR-L858R/T790M) were performed with the HTRF KinEASE-TK assay from Cisbio according to the manufacturer's instructions. Accordingly, the amount of EGFR in each reaction well was set to 0.60 ng EGFR-wt (0.67 nM), 0.10 ng EGFR-L858R (0.11 nM), 0.07 ng EGFR-T790M/L858R (0.08 nM) or 0.80 ng EGFR-L858R/T790M/C797S (0.88 ng), respectively. An artificial substrate peptide (TK-substrate from Cisbio) was then phosphorylated by EGFR. After completion of the reaction (reaction times: 25 min for wt, 15 min for L858R, 20 min for T790M/L858R and 10 min for L858R/T790M/C797S), the reaction was stopped by the addition of a buffer containing EDTA, an anti-phosphotyrosine antibody labeled with europium cryptate and streptavidin labelled with the fluorophore XL665. Fluorescence resonance energy transfer (FRET) between europium cryptate and XL665 was then measured after an additional hour of incubation to quantify the phosphorylation of the substrate peptide. ATP concentrations were set at their respective K_m-values (9.5 μM for wt, 9 μM for L858R, 4 μM for L858R/T790M and 11 μM for L858R/T790M/C797S) while a substrate concentration of 1 μM, 225 nM, 200 nM and 325 nM, respectively, were used.

Chapter 3 – Overview of Aims, Objectives and General Practices

Kinase and inhibitor were preincubated for 30 min before the reaction was started by addition of ATP and substrate peptide. An EnVision multimode plate reader (Perkin Elmer) was used to measure the fluorescence of the samples at $\lambda = 620$ nm (Eu³⁺-labeled antibody) and $\lambda = 665$ nm (XL665 labeled streptavidin) 50 μ s after excitation at $\lambda = 320$ nm. The quotient of both intensities for reactions made with eight different inhibitor concentrations was then analyzed using the Quattro Software Suite for IC₅₀ determination. Each reaction was performed in duplicate and at least three independent determinations of each IC₅₀ value were made.

3.7.2.4 Cell Culture

A431, HCC827, H1975 and A549 cells were obtained from the American Type Culture Collection (ATCC). A431 cells were cultured in DMEM high glucose media (Life Technologies), and HCC827, H1975 and A549 cells were cultured in RPMI media (Life Technologies). Both contained L-glutamine and were supplemented with 10% FBS (GE Healthcare) and 1% PenStrep (Life Technologies) in a humidified incubator at 37 °C with a 5% CO₂ atmosphere. Cell line authentication was been performed last in August 2018 by STR profiling of 16 alleles.

3.7.2.5 Viability Assay (EC₅₀ Determination)

Cells were seeded at cell numbers that insured linearity and optimal signal intensity (150 – 300 cells/well, 25 μ L) and cultured for 24 hours in serum- and antibiotic-containing media in a humidified chamber at 37 °C with a 5% CO₂ atmosphere. Cells were then treated with EGFR inhibitors in serial dilutions (14 nM to 30 μ M) in DMSO, with Staurosporine as control, and incubated for 96 hours. Afterwards viability studies were carried out using the CellTiter-Glo Assay (Promega), a homogeneous method of determining the number of viable cells in culture. Based on quantification of ATP, the assay indicates the presence of metabolically active cells. For cellular evaluation, the CellTiter-Glo reagent was prepared according to the instructions of the kit and diluted in 1:1 ratio with the complete growth medium suitable for the corresponding cell line. Thereafter, the reagent and assay plates were equilibrated at room temperature for 20 min, with equal volumes of the reagent being added to the volume of culture medium present in each well (25 μ L). The microplates were mixed for 2 minutes on an orbital shaker to induce cell lysis and were then incubated at room temperature for 20 minutes for stabilization of the luminescent signal. Following incubation, the luminescence was recorded on an EnVision microplate reader (Perkin Elmer) using 500 ms integration time. The data was then analysed using the Quattro Research Software Suite for EC₅₀ determination. As quality control, the Z'-factor was calculated from 16 positive and negative control values. Only assay results showing a Z'-factor ≥ 0.5 were used for further analysis. All data points were measured in triplicate for each plate and were replicated in at least two plates.

3.7.2.6 Covalent Mass Spectrometry Experiments

The drug-resistant EGFR-T790M mutant variant was utilised for covalent mass spectrometry experiments. A mixture of 26.4 μ M of protein was incubated with 79.2 μ M of inhibitor in buffer (25 mM TRIS, 250 mM NaCl, 10% glycerol, 1 mM TCEP, pH 8) on ice for 1 h. Aliquots were analysed by ESI-MS using a Thermo Fisher Scientific Dionex UltiMate 3000 HPLC system connected to a Thermo Fisher Scientific Velos Pro (2D ion trap).

Chapter 3 – Overview of Aims, Objectives and General Practices

Accordingly, 5 μL of sample was injected and separated using a Vydac 214TP C4 5 μm column (150 mm x 2.1 mm) with two mobile phases which included A (0.1% TFA in water) and B (acetonitrile). Separation was initiated with 20% solvent B for 5 min, followed by a gradient increase to 90% solvent B over 14 min at a flow rate of 210 $\mu\text{L}/\text{min}$. A mass range of 700 to 2000 m/z was scanned, and raw data was deconvoluted and analysed using MagTran software.¹¹

3.8 References

1. T. Ishikawa, M. Seto, H. Banno, Y. Kawakita, M. Oorui, T. Taniguchi, Y. Ohta, T. Tamura, A. Nakayama, H. Miki, H. Kamiguchi, T. Tanaka, N. Habuka, S. Sogabe, J. Yano, K. Aertgeerts and K. Kamiyama, *Journal of Medicinal Chemistry*, 2011, **54**, 8030-8050.
2. J. Beenstock, N. Mooshayef and D. Engelberg, *Trends in Biochemical Sciences*, 2016, **41**, 938-953.
3. G. H. Iyer, S. Garrod, V. L. Woods and S. S. Taylor, *Journal of Molecular Biology*, 2005, **351**, 1110-1122.
4. R. Dasari, A. De Carvalho, D. C. Medellin, K. N. Middleton, F. Hague, M. N. M. Volmar, L. V. Frolova, M. F. Rossato, J. J. De La Chapa, N. F. Dybdal-Hargreaves, A. Pillai, V. Mathieu, S. Rogelj, C. B. Gonzales, J. B. Calixto, A. Evidente, M. Gautier, G. Munirathinam, R. Glass, P. Burth, S. C. Pelly, W. A. L. van Otterlo, R. Kiss and A. Kornienko, *ChemMedChem*, 2015, **10**, 2014-2026.
5. Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame and J. Taunton, *Journal of the American Chemical Society*, 2017, **139**, 680-685.
6. H. Zhang, W. Wu, C. Feng, Z. Liu, E. Bai, X. Wang, M. Lei, H. Cheng, H. Feng, J. Shi, J. Wang, Z. Zhang, T. Jin, S. Chen, S. Hu and Y. Zhu, *European Journal of Medicinal Chemistry*, 2017, **135**, 12-23.
7. J. Jang, J. B. Son, C. To, M. Bahcall, S. Y. Kim, S. Y. Kang, M. Mushajiang, Y. Lee, P. A. Jänne, H. G. Choi and N. S. Gray, *European Journal of Medicinal Chemistry*, 2017, **136**, 497-510.
8. J. Pettinger, K. Jones and M. D. Cheeseman, *Angewandte Chemie International Edition*, 2017, **56**, 15200-15209.
9. K. S. Thress, C. P. Paweletz, E. Felip, B. C. Cho, D. Stetson, B. Dougherty, Z. Lai, A. Markovets, A. Vivancos, Y. Kuang, D. Ercan, S. E. Matthews, M. Cantarini, J. C. Barrett, P. A. Jänne and G. R. Oxnard, *Nature Medicine*, 2015, **21**, 1-5.
10. W. L. F. Armarego and C. L. L. Chai, *Purification of Laboratory Chemicals*, Butterworth-Heinemann 7th edn., 2013.
11. Z. Zhang and A. G. Marshall, *Journal of the American Society for Mass Spectrometry*, 1998, **9**, 225-233.

Chapter 4

Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Abstract

The body of work discussed in this chapter comprises of the design, synthesis and biochemical and cellular evaluation of a combinatorial library of compounds based on a previously synthesised lead candidate. The identification and synthesis of this lead compound will be alluded to, followed by a report of the dual EGFR/HER2 inhibitors forming the molecular inspiration of this research. Derivatisation with various glycol chain moieties in key substitution positions, in the absence or presence of a solubilising group and/or an electrophilic Michael acceptor, resulted in the synthesis of a broad combinatorial library. This library consists of series A - H, containing 26 reversible and potentially irreversible inhibitors. The synthesis of these compounds is discussed in full, with emphasis on reactions optimized during the generation of these inhibitors. The synthesised molecules were then screened against numerous mutant variants of the EGFR enzyme, both biochemically and in a cellular context, revealing specific trends affecting the activity and binding mode of the compounds. While picomolar range IC₅₀ values were observed for biochemical inhibition, these values were unfortunately not reflected against corresponding cell lines. Evaluation against HER2 mutant cell lines exhibited the compounds potential for development as dual EGFR/HER2 inhibitors. Lastly, a crystal structure of a potent inhibitor in the cSrc-T338M/S345C surrogate, reinforced our postulates regarding the observed trends in activity.

4.1 Introduction

The work undertaken in this chapter was completed during a four-month research visit to Prof. Dr. Daniel Rauh's laboratories at the Fakultät für Chemie und Chemische Biologie, Technische Universität Dortmund, Germany. This visit was made possible by funding from the South African Medical Research Council (MRC) and National Research Foundation (NRF). This body of research focusses on the optimization of a pyrrolopyrimidine-based lead compound previously identified and synthesized by Mr. Philip Klövekorn, a student under the supervision of Prof. Rauh. We will briefly discuss the identification and optimization towards this lead compound, followed by the molecular musings and hypothesis for this body of research. This encompasses the derivatisation and interchanging of specific structural elements on the lead compound scaffold, providing valuable information correlating to binding characteristics and interactions of the inhibitor within the ATP binding cleft.

4.2 Lead Compound Identification and Optimization

An analysis of ligands bound to the EGFR kinase domain within the protein data bank, led the Rauh research group to the pyrrolopyrimidine-based compound **AEE788**, shown in **Figure 4.1a**. AEE788 is a dual family EGFR/HER2 and VEGFR tyrosine kinase inhibitor, with antitumor and antiangiogenic properties, which was developed by Novartis in 2004.¹ As a 1st generation inhibitor, the compound displayed efficient inhibition of the EGFR-bearing activating mutations such as EGFR-L858R. However, as with other 1st generation EGFR inhibitors such as gefitinib and erlotinib, the onset of the T790M gatekeeper mutation resulted in a drastic diminishment in efficacy. This decrease in potency could be attributed to the bulky mutated methionine residue creating a steric clash with the 1-phenylethylamine moiety of AEE788, as illustrated in **Figure 4.1b**, leading to an altered binding orientation. Further complications associated with unacceptable toxicity resulted in clinical trials of the reversible inhibitor being discontinued in 2006.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

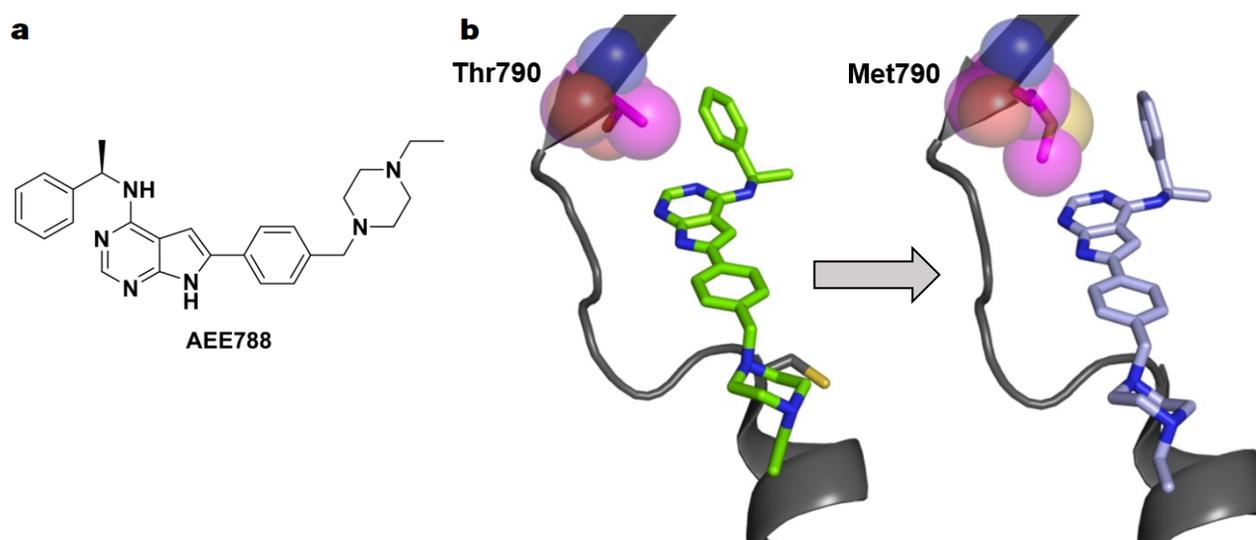


Figure 4.1: a) Structure of AEE788 and b) Steric effects of EGFR-T790M on the binding mode of AEE788 (PDB:2J6M and 2JIU).

To circumvent the pitfalls associated with the T790M mutation, the Rauh group set about optimizing the pyrrolopyrimidine scaffold, with their strategy illustrated in **Figure 4.2a**. Firstly, the 1-phenylethylamine group was substituted with a methoxy group at the 4-position, shown to be small enough to avoid negative steric interference with the Met797 gatekeeper side chain. A variety of solubilizing groups were then assessed, with a methyl-piperazine moiety linked directly to the aromatic ring being found to be the most efficacious. Lastly, supported by structural analysis, the 5-position of the pyrrolopyrimidine scaffold was found to be most suitable for the introduction of a phenyl linker bearing a reactive Michael acceptor. Utilization of an acrylamide moiety in the *meta*-position presented a suitable binding orientation and adjacent proximity to Cys797, allowing the warhead to potentially undergo covalent bond formation. For their efforts, the group were rewarded with lead compound **1**, whose binding mode can be seen in **Figure 4.2b**. The pyrrolopyrimidine core was seen to form bidentate hydrogen bond interactions with Met793 of the kinase hinge region, with the solubilising group directly exposed to the solvent. The favourable positioning of the acrylamide relative to Cys797 can also be seen below.

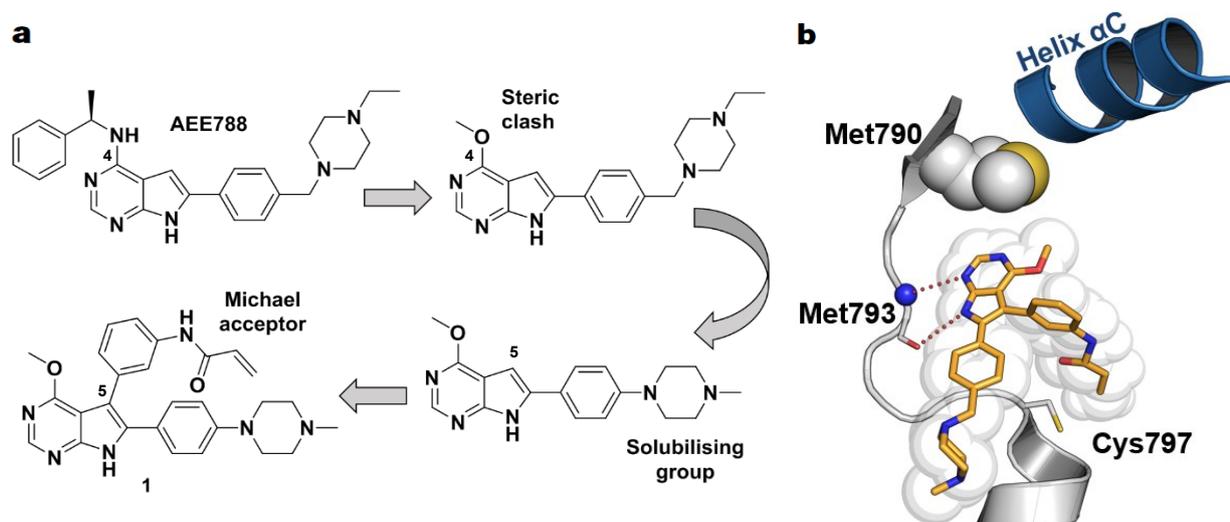


Figure 4.2: a) Synthetic optimization of AEE788, affording lead compound **1** and b) Binding mode of compound **1**, illustrating hydrogen bond interactions with the hinge region, the solubilising group and proximity for covalent inhibition (PDB: 2JIT).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

The lead compound was tested in an activity-based, EGFR HTRF biochemical assay and was shown to inhibit all variants of EGFR kinases with sub-nanomolar IC_{50} values comparable to that of osimertinib. As can be seen below in **Figure 4.3**, compound **1** inhibited the EGFR-wt kinase with an IC_{50} of 0.41 nM, exhibiting a 4- and almost 2-fold selectivity for the L858R (0.12 nM) and L858R/T790M (0.24 nM) mutant variants respectively. Moreover, the reversible counterpart **2**, bearing a propionamide moiety in place of the Michael acceptor, showed a marked decrease in efficacy, with IC_{50} value of 305 nM against the L858R/T790M double mutant. This equated to over a thousand-fold loss in activity, substantiating the formation of a covalent bond and highlighting the importance of an irreversible mode of inhibition in targeting the double mutant.

Similar trends were observed in subsequent cellular tests of the compounds. A wild-type bearing cell line A431, EGFR-delE746_A750 activating mutation-carrying HCC827 cells, the EGFR-L858R/T790M 2nd generation drug resistant cell line H1975 and the clinically relevant EGFR-delE746_A750/T790M/C797S triple mutant cell line were employed in this assay. While compound **1** displayed slightly more favourable activity than osimertinib in a biochemical setting, the lead compound was found to be nearly 4-fold less effective against the H1975 cell line, with an EC_{50} of 60 nM compared to the blockbuster drug (16 nM).

Examination of the biochemical results against the L858R/T790M/C797S triple drug resistant mutant revealed that compound **1** (50 nM) was found to be more than twice as potent as osimertinib (113 nM). However, when compared to the L858R/T790M double mutant, both compounds suffered a 200-fold loss of activity, reflecting the current crisis in clinical resistance to 3rd generation inhibitors. This sentiment is mirrored in the cellular activities of the compounds against the PC9/T790M/C797S triple mutant. Forthwith, we strove to improve the biochemical and cellular efficacy of compound **1** against the double and triple mutant by creating favourable reversible interactions within the active site of the kinase and improving on pharmacokinetic parameters such as metabolic stability and bioavailability. The results of this endeavour will be described within this chapter.

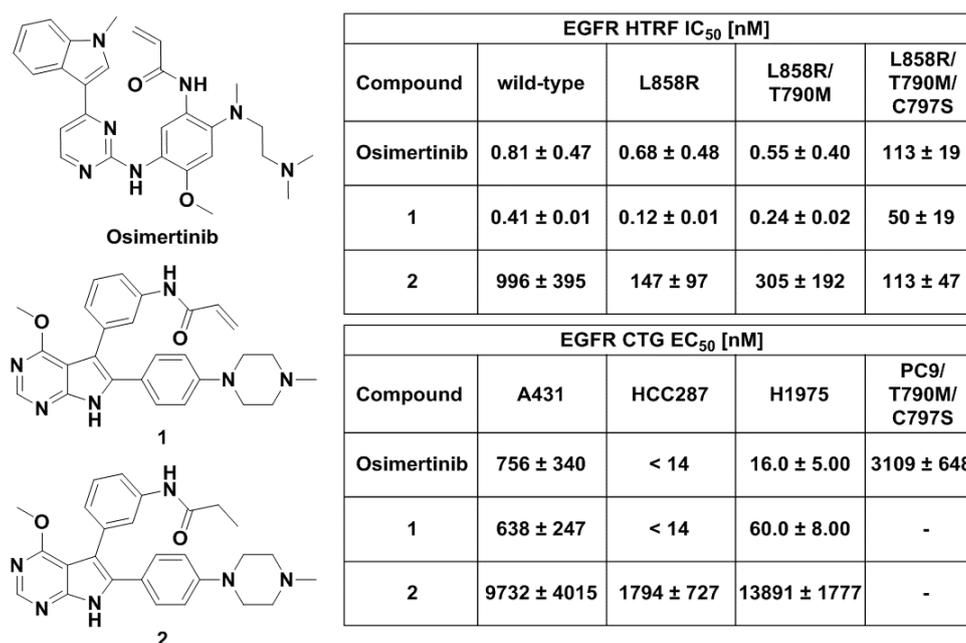


Figure 4.3: Biochemical and cellular assay activity of osimertinib, lead compound **1** and its reversible counterpart **2**.

4.3 Inspiration from Dual EGFR/HER2 Inhibitors

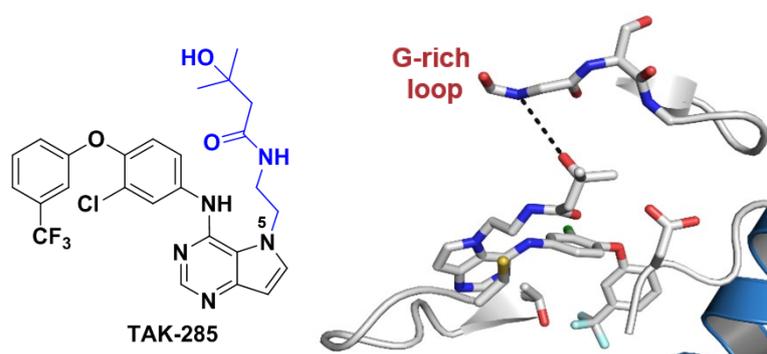


Figure 4.4: Structure of the dual family inhibitor TAK-285 with binding mode highlighting interactions with the glycine-rich loop.

Tak-285, pictured above in **Figure 4.4**, is a dual ErbB inhibitor of EGFR and HER2 that has been investigated for breast, lung, gastric and prostate cancer.² The pyrrolopyrimidine compound was designed to bind reversibly within the ATP cleft, displaying potent inhibition of wild-type and activating mutation bearing EGFR, HER2 and anti-tumour growth activity. The molecule exploits the hydrophobic pocket located behind the mutated Met790 residue, formed by side chains near the DFG motif, with a trifluoromethyl phenoxy back-pocket binding element. As Kamiyama and co-workers made use of 5*H*-pyrrolo-[3,2-*d*]pyrimidine as their heterocyclic scaffold, instead of the 7*H*-pyrrolo-[2,3-*d*]pyrimidine used for lead compound **1**, the driving group only formed a single hydrogen bond interaction with the Met793 residue in EGFR.³ This resulted in a nonoptimal hinge region hydrogen bonding pattern and consequent weak inhibitory activity against the L858R/T790M mutant. However, use of the 5*H*-pyrrolo-pyrimidine scaffold allowed the group to introduce various substituents at the *N*-5 position, such as the carboxamide moiety (**Figure 4.4**, blue) utilised for TAK-285. The terminal hydroxy moiety of long alkyl chain substituents were shown to form hydrogen bond interactions with the highly conserved glycine-rich loop (**Figure 4.4**, red), enhancing the reversible binding affinity of the inhibitor.² Furthermore, derivatives bearing substituents on the *N*-5 position of the scaffold exhibited significantly more potent cellular growth inhibitory activity and improved metabolic stability.²

One such derivative is the compound **SYR127063**, shown in **Figure 4.5a** on the following page, which retains the heterocyclic core and hydrophobic back pocket binding element of TAK-285. SYR127063 is structurally identical to TAK-285, with the exception of a 2-(2-hydroxyethoxy)ethyl chain substituent at the *N*-5 position (**Figure 4.5a**, blue). Incorporation of this glycol chain gave rise to favourable hydrogen bond interactions with the glycine-rich loop (**Figure 4.5a**, red) leading to superior biochemical inhibition of both EGFR and HER2 over TAK-285. Moreover, SYR127063 showed similar tumour cell growth inhibition activity, *in vivo* efficacy in tumour xenograft mouse models without body loss and excellent metabolic stability in mouse and human studies.²

The desirable cellular activity, metabolic stability, pharmacokinetic profiles and *in vivo* antitumor efficacy of these two inhibitors proved to be a source of inspiration for the optimization of compound **1**. It was postulated that incorporation of the 2-(2-hydroxyethoxy)ethyl of SYR127063 within the lead compound scaffold could be beneficial and lead to an improvement in the traits mentioned above. Through structural analysis we set out to determine the best positions on compound **1** to integrate the glycol chain and effectuate interactions with the glycine-rich loop.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

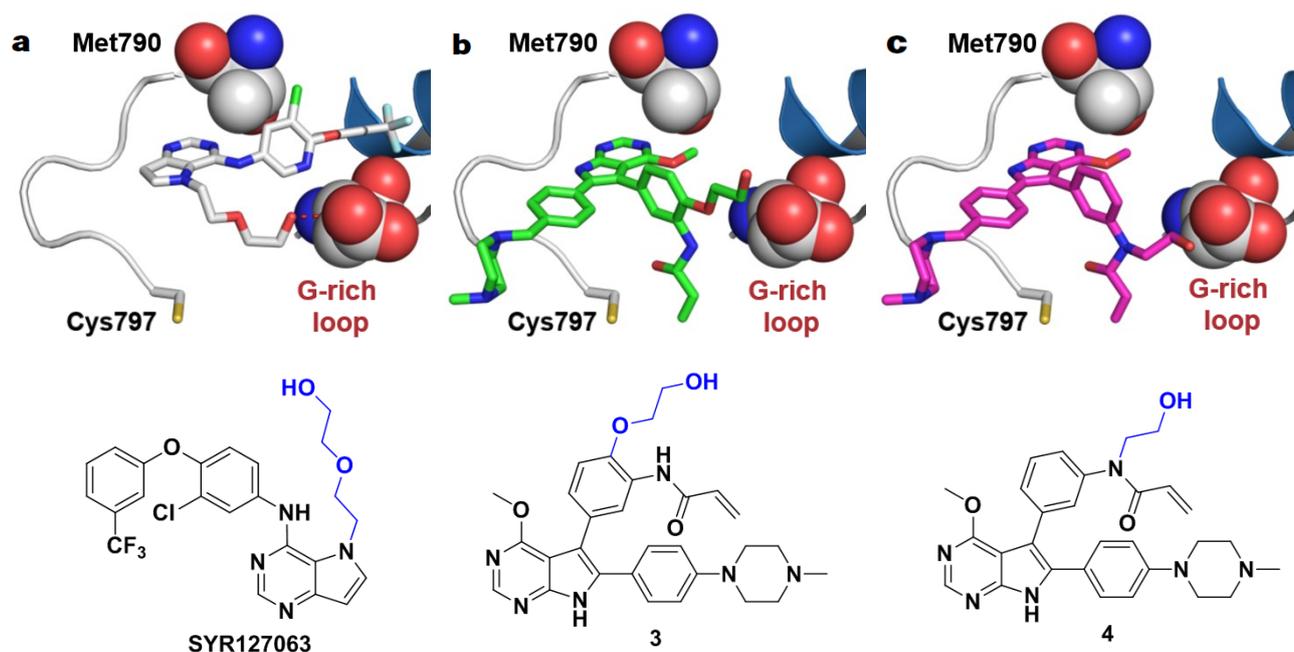


Figure 4.5: Structure and binding mode of a) TAK-285-derived compound SYR127063, b) Compound **3** with 2-hydroxyethoxy chain in an *ortho*-position to the acrylamide and c) Compound **4** with 2-hydroxyethyl chain at the *N*-position of the acrylamide (PDB: 3PPO).

Through molecular modelling, it was found that incorporation of a 2-hydroxyethoxy motif in an *ortho*-position to the acrylamide (**Figure 4.5b**, compound **3**, blue), would provide the highest likelihood to form interactions with the glycine-rich loop (**Figure 4.5b**, red) and enhance reversible binding affinity. The two-carbon spacer glycol chain was found to be optimal, able to avoid a steric clash and interact with the glycine-rich loop, while not interfering with the binding mode of the inhibitor. This would preserve the critically important hinge region interactions with Met793 and covalent bond formation with Cys797. Similarly, inclusion of a 2-hydroxyethyl moiety directly onto the acrylamide (**Figure 4.5c**, compound **4**, blue) was found to be an interesting prospect. Substitution on the *N*-position would not only provide the beneficial interactions proposed with compound **3**, but potentially improve the warhead's metabolic stability and reactivity towards covalent bond formation.^{4, 5} We therefore embarked on planning the synthesis of a library of lead-derived compounds bearing glycol chains at these desired positions.

4.4 Objectives and Target Synthetic Combinatorial Library

With the optimal positioning and length of the glycol chain substituents determined, we elected to make use of three variants based on the chain utilised in SYR12706. However, contrary to the terminal hydroxy functionality employed in this inhibitor, we chose to make use of a terminal ether functionality. We hypothesized that the additional electron donating effect, and thus electron density on the oxygen atom of the ether, would serve to form more favourable hydrogen bond interactions with the protonated glycine residues.⁶ The structures of these groups are highlighted in blue in **Figure 4.6** and are namely 2-methoxyethyl, 1-methoxypropan-2-yl or 2-ethoxyethyl for the *N*-substituted position (**Figure 4.6a**, blue) and 2-methoxyethoxy, (1-methoxypropan-2-yl)oxy and 2-ethoxyethoxy for the *ortho*-position (**Figure 4.6b**, blue). Our efforts were therefore focussed on synthesising a combinatorial library of 24 compounds, where the investigation outputs were centred around three parameters.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

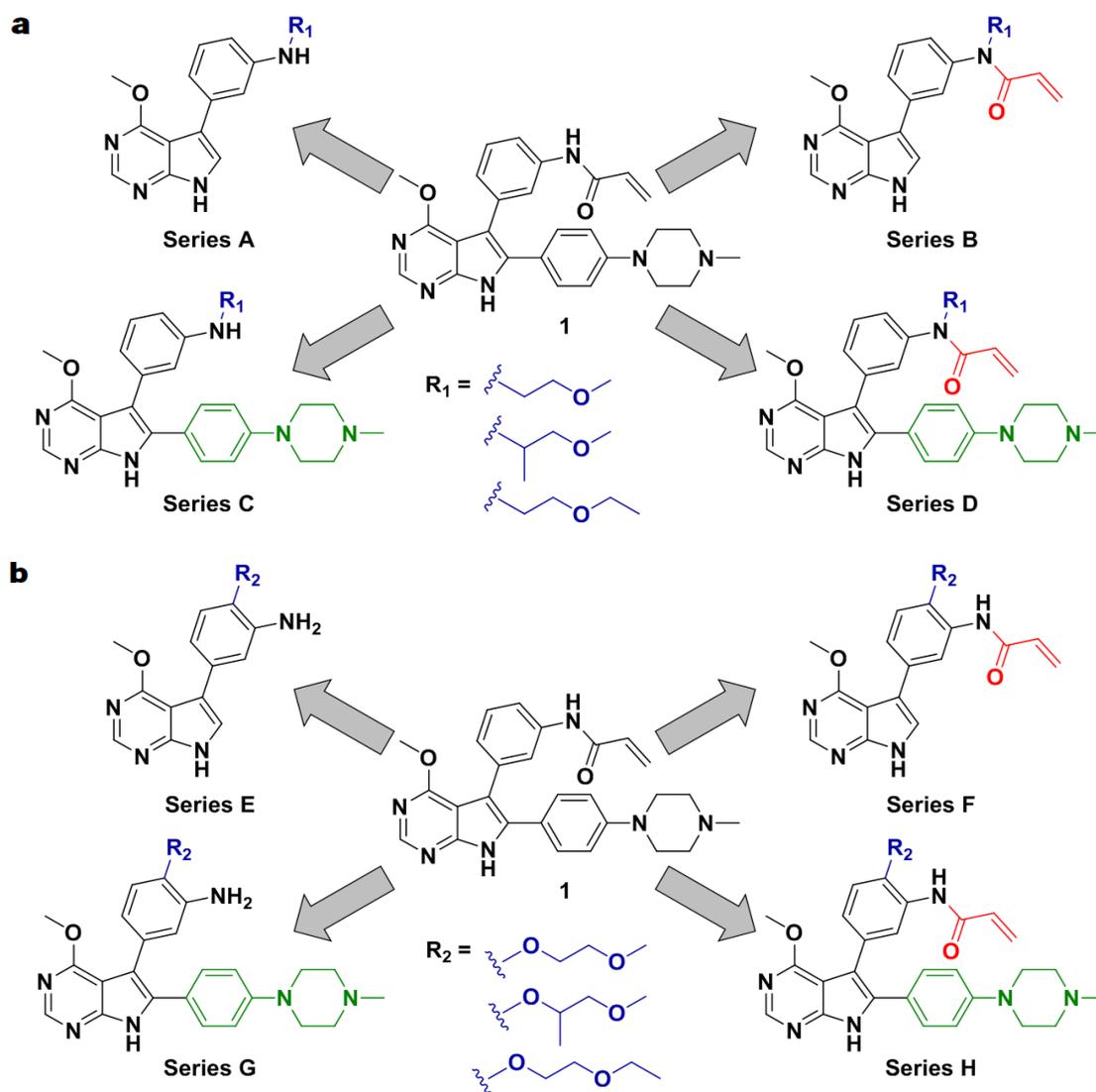


Figure 4.6: Combinatorial synthetic library containing a) Series A – D, with the glycol chain installed on the *N*-position of the acrylamide and b) Series E – H, with the glycol chain in an *ortho*-position to the acrylamide.

Firstly, effects of substitution of the glycol chains at either the *N*- or *ortho*-position of the acrylamide would be explored. We envisaged that the addition of these flexible groups would increase binding affinity by the formation of new reversible interaction with the glycine-rich loop, as discussed previously. The optimal total chain length and substitution pattern could thus be determined by the gains or losses in biochemical and cellular efficacy. Secondly, analogues of these compounds with and without the electrophilic Michael acceptor would be synthesised to specifically determine their potential as irreversible inhibitors of the T790M double mutant and as reversible inhibitors of the C797S triple mutant respectively. Lastly, the necessity of the phenyl methyl-piperazine solubilising group would be considered. The effects of its presence or absence on the scaffold, particularly with regards to cellular efficacy, would be scrutinised, indicating whether the addition of the glycol chain could be sufficient to replace it.

From these considerations, a summary of the compounds to be synthesised can be seen in **Figure 4.6** above. **Figure 4.6a** encompasses all compounds bearing the glycol chain moiety at the *N*-position of the acrylamide, while **Figure 4.6b** displays those at the *ortho*-position. From **Figure 4.6a**, series A and B would both bear the glycol chains (blue) without the solubilising group, and without and with the Michael acceptor (red), respectively.

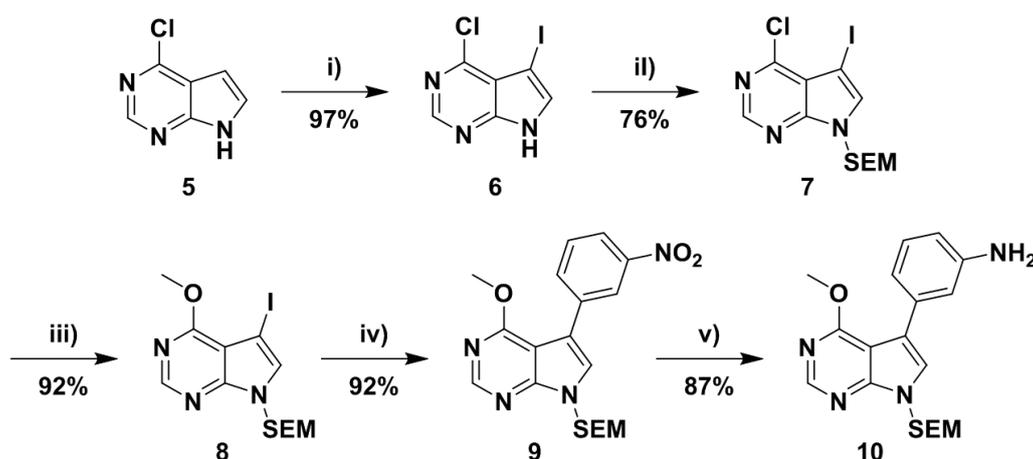
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Series C and D follow the same trend but with the presence of the solubilising group (green) for both. The conditions described were applied for series E – H in **Figure 4.6b**, with the exception of the glycol chains on the *ortho*-position of the acrylamide for this library. With the target combinatorial library and objectives outlined, we embarked on the synthesis of the compounds comprising series A – H.

4.5 Combinatorial Library Synthesis

4.5.1 Pyrrolopyrimidine Scaffold Synthesis

The synthesis of the pyrrolopyrimidine core scaffold **10**, shown in **Scheme 4.1**, had been previously completed in the synthesis of lead compound **1**. However, several adjustments in reagents and reactions conditions implemented during the synthesis of this library of compounds, resulted in a fully optimized synthetic route. These modifications will be highlighted in the synthetic discussion that follows. Starting from the commercially available 4-chloropyrrolopyrimidine **5**, installation of an iodo functionality at the 5-position was undertaken using *N*-iodosuccinimide. While initial procedures employed acetonitrile as solvent, use of DMF in the absence of light provided a near quantitative yield of the iodinated compound **6**, which was isolated after workup with no further purification required. Owing to its stability and frequent use in the protection of *N*-heterocycles, the 2-trimethylsilylethoxymethyl (SEM) group was selected as protecting group for the synthesis. Following deprotonation of the pyrrole nitrogen at the 7-position using sodium hydride in DMF, the SEM-group was attached affording compound **7** in a yield of 76%. This allowed us to introduce the sterically innocuous methoxy group at the electrophilic 4-position of the pyrimidine ring. While harsh conditions involving use of sodium methoxide were originally imposed, we found that simply stirring compound **7** in methanol with K_2CO_3 afforded compound **8** in excellent yields. Next, optimization of the Suzuki-Miyaura coupling was explicitly investigated, as this reaction would be required in other instances throughout the syntheses.



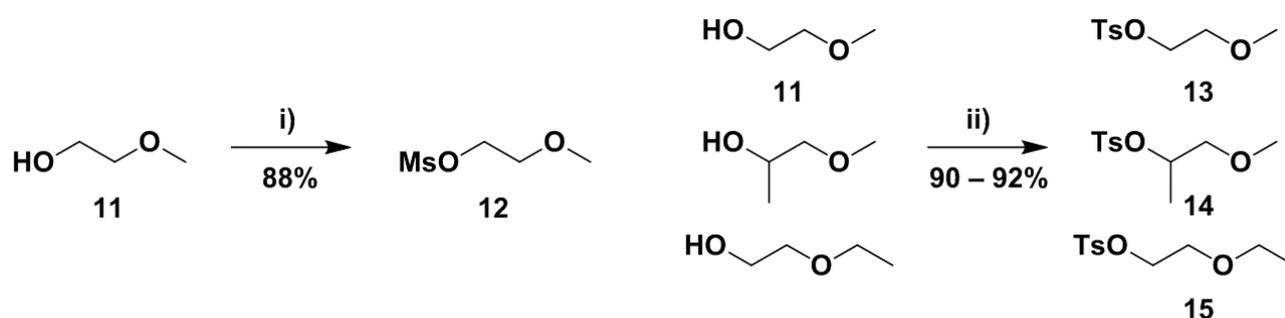
Scheme 4.1: Five step synthesis of pyrrolopyrimidine core scaffold **10**. Reagents and conditions: i) NIS (1.05 equiv.), DMF, rt, 16 h; ii) NaH (1.17 equiv.), SEM-Cl (1.23 equiv.), DMF, $-10\text{ }^{\circ}\text{C}$ – rt, 16 h; iii) K_2CO_3 (2.2 equiv.), MeOH, rt, 16h; iv) (3-nitrophenyl)boronic acid (1.2 equiv.), $Pd(PPh_3)_4$ (5 mol%), Na_2CO_3 (2.5 equiv.), dioxane/ H_2O (3:1), $80\text{ }^{\circ}\text{C}$, 16 h; v) Fe (5 equiv.), NH_4Cl (1 equiv.), EtOH/ H_2O (1:1), ultrasonication at $60\text{ }^{\circ}\text{C}$, 6 h.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Use of multiple solvents, including toluene, ethanol, dioxane and water, in combination with either sodium or potassium carbonate, revealed the optimal reaction conditions for the coupling. We found that a 3:1 mixture of dioxane/water, a catalyst load of 5 mol% tetrakis(triphenylphosphine)palladium(0) and sodium carbonate serving as base, in a sealed schlenk at 80 °C, consistently provided yields above 90%. Furthermore, use of the freeze-pump-thaw technique, typically requiring 3 – 5 cycles to sufficiently degas the solvent mixture prior to reaction, was found to be essential for complete consumption of the starting material **8**. This provided the 3-nitrophenyl bearing pyrrolopyrimidine compound **9** in a yield of 92% and optimized reaction conditions for further application. Lastly, reduction of the aromatic nitro group was investigated using either Pd/C with ammonium formate/H₂ gas or iron and ammonium chloride as reducing agents, while heated under reflux or ultrasonic agitation. Use of iron and ammonium chloride in a 1:1 mixture of ethanol/water, while undergoing ultrasonication at 60 °C, was found to be most robust and high yielding, affording compound **10** in an 86% yield after purification. This provided an economical and reliable route to the core scaffold, which could be carried out on a large scale of up to 20 g, allowing us to pursue the synthesis of the combinatorial library.

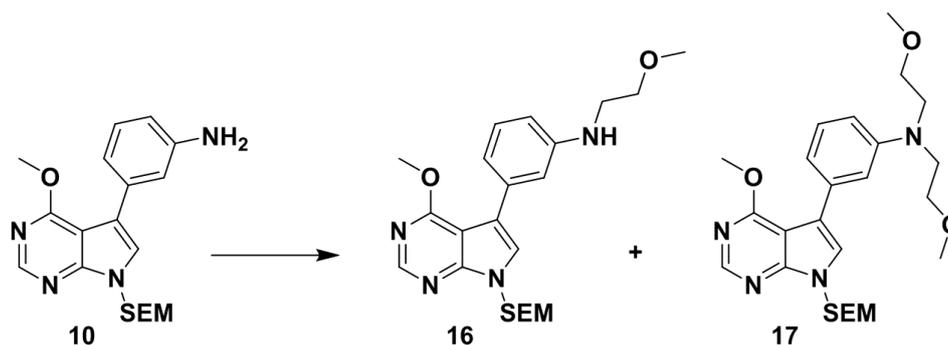
4.5.2 Synthesis of Series A and B

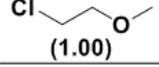
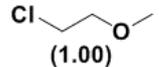
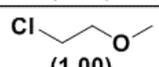
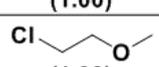
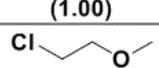
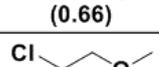
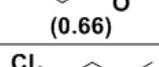
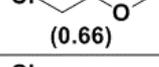
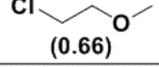
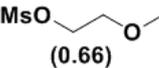
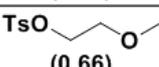
The first step in the synthesis of series A was the monoalkylation of the newly formed pyrrolopyrimidine-aniline structure **10** with our chosen glycol chain moieties. Initial attempts met with the sole formation of the undesired dialkylated product, necessitating optimization of reactions conditions and reagents to selectively afford the monoalkylated product. For this optimization, the 2-methoxyethyl chain was chosen for its structural simplicity, while exploring varying combinations of temperatures and heat sources, bases, solvents and leaving groups. Synthesis of the mesylated and tosylated 2-methoxyethyl fragments is shown below in **Scheme 4.2**. From the commercially available **11**, reaction with either mesyl- or tosyl-chloride and Et₃N serving as base in DCM, furnished compound **12** and **13** respectively in excellent yields. Furthermore, once it was found that the tosyl-leaving group functionality provided the highest overall selectivity and yield in the monoalkylation reaction, the tosylated compounds **14** and **15** were synthesized in a similar manner, for future use.



Scheme 4.2: Synthesis of glycol ether chains **12** and **13** – **15** with mesyl- and tosyl-leaving groups respectively. Reagents and conditions: i) methanesulfonyl chloride (1.1 equiv.), Et₃N (1.2 equiv.), DCM, –10 °C – rt, 16 h; ii) *p*-toluenesulfonyl chloride (1.1 equiv.), Et₃N (1.2 equiv.), DCM, –10 °C – rt, 16 h.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Table 4.1: Optimization table for the monoalkylation of pyrrolopyrimidine scaffold **10**.

Monoalkylation Optimization Table									
Entry	Leaving group (Equiv.)	Time (h)	Solvent	Heat (Source)	Base (Equiv.)	Additive (Equiv.)	Monoalkylated product 16 (%)	Dialkylated product 17 (%)	Yield BRSM (%)
1	 (1.00)	0.5	DMF	110 °C (MW)	K ₂ CO ₃ (3.00)	KI (1.00)	10	39	-
2	 (1.00)	0.5	DMF	110 °C (MW)	DIPEA (3.00)	KI (1.00)	6	45	-
3	 (1.00)	16	DMF	rt	NaH (2.00)	-	-	-	-
4	 (1.00)	16	DMF	100 °C	K ₂ CO ₃ (3.00)	KI (1.00)	12	31	-
5	 (0.66)	16	DMF	100 °C	K ₂ CO ₃ (3.00)	KI (1.00)	30	8	61
6	 (0.66)	16	DMF	80 °C	K ₂ CO ₃ (3.00)	KI (1.00)	35	-	67
5	 (0.66)	16	DMF	80 °C	K ₂ CO ₃ (1.50)	KI (1.00)	41	-	69
7	 (0.66)	16	ACN	80 °C	K ₂ CO ₃ (1.50)	KI (1.00)	33	-	67
8	 (0.66)	16	DMF	80 °C	K ₂ CO ₃ (1.50)	-	27	-	58
9	 (0.66)	16	DMF	80 °C	K ₂ CO ₃ (1.50)	-	60	-	82
10	 (0.66)	32	DMF	80 °C	K ₂ CO ₃ (1.50)	-	65	-	90

Optimization results for monoalkylation of fragment **10** is tabulated above in **Table 4.1**, with the mono- and dialkylated products **16** and **17** shown at the top. Early attempts using 2-chloroethyl methyl ether as electrophile with potassium carbonate or DIPEA as base in DMF at 110 °C under microwave irradiation afforded **17** as the major product with trace amounts of **16** (**Entries 1** and **2**). This included the use of potassium iodide to promote Finkelstein halogen exchange to potentially increase the reactivity of the electrophile. Use of sodium hydride at ambient temperature in **Entry 3** resulted in the return of the starting material **10**, with no reaction having taken place. We thus abandoned the use of microwave irradiation and pursued potassium carbonate as the base of choice for subsequent reactions. However, use of conventional heating at 100 °C in **Entry 4** displayed a similar result to that of **Entry 1**, with only 12% of **16** being isolated.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

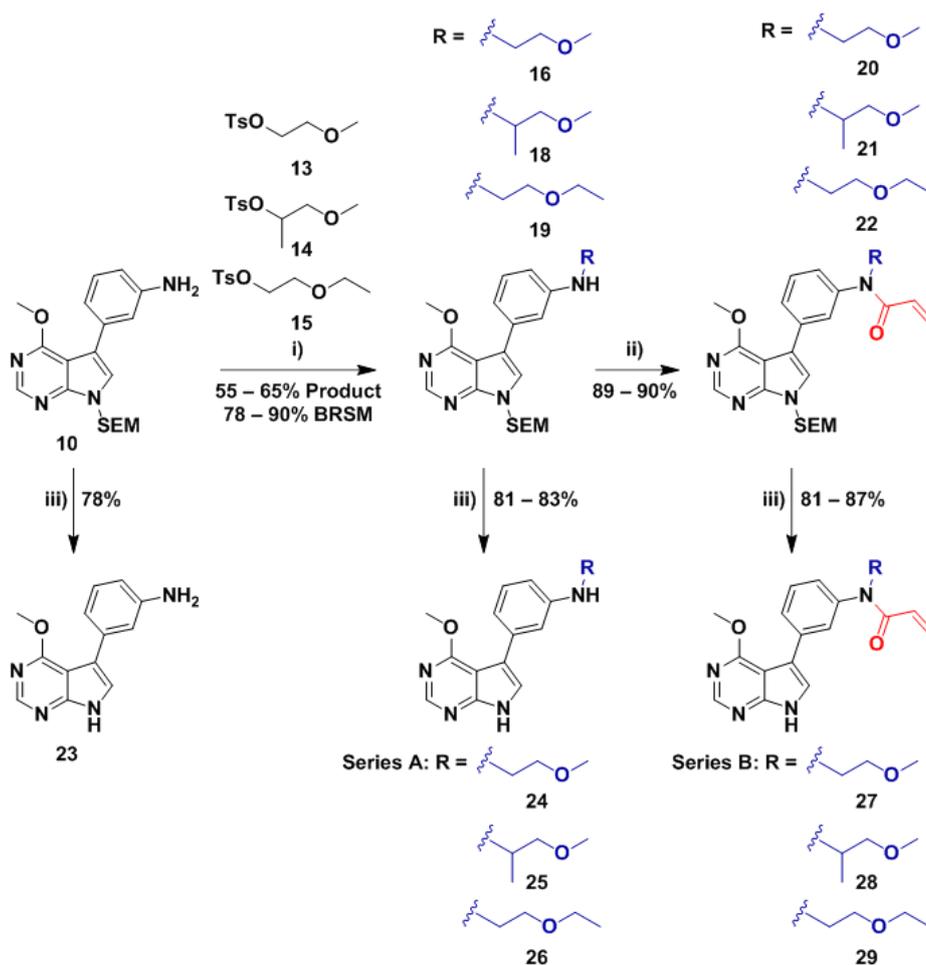
We concluded that formation of the dialkylated product is favoured, owing to the inductive effect of the glycol chain. This increases the acidity of the amine and thereby the nucleophilicity of fragment **16** over the starting material **10**. Consequently, this created a higher likelihood of a second nucleophilic attack by **16** on the glycol electrophile, resulting in the dialkylated **17** being driven over formation of **16**. We therefore theorized that reducing the available amount of the alkyl halide might lessen the possibility for and occurrence of this second nucleophilic attack. Gratifyingly, when we reduced the 2-chloroethyl methyl ether used in the reaction to 0.66 equivalents (**Entry 5**), we saw a reversal in the product distribution, isolating 30% of **16** and a trace amount of **17**. Moreover, as the leaving group moiety was the limiting reagent, we were able to isolate the excess starting compound **10** and recycle the recovered portions in successive reactions and calculate a yield base on the recovery of the starting material (BRSM).

To further minimize the potential for a second nucleophilic attack and formation of **17**, we also lowered the reaction temperature to 80 °C and decreased the amount of base used in **Entries 6** and **7** respectively. This further reinforced our hypothesis, as the cumulative effects of lower temperature and base resulted in isolation of up to 41% of **16** with a BRSM yield of 69%. Interchange of DMF with acetonitrile as solvent in **Entry 8** resulted in minor loss in the yield of compound **16**, potentially ascribed to the slightly favoured attributes of DMF in promoting the S_N2 reaction rate (lower dielectric constant and polarity).⁷

Lastly, we proceeded to investigate the effect of a mesyl- or tosyl-leaving group on the mono- and dialkylated product substitution pattern. To our delight, reaction with the mesylated and tosylated 2-methoxyethyl glycol chain, shown in **Entries 8** and **9**, resulted in isolation of 33% and 60% of compound **16** and a BRSM yield of 63% and 82% respectively. Beneficially, the tosyl-functionality allowed for simple UV visualization on TLC and enhanced awareness of electrophile consumption rates and reaction completion time. Encouraged by this result, we continued with the tosylated glycol chain and doubled the reaction time in **Entry 10**, eliciting a 5% increase in the isolated yield to 65% and a BRSM yield of 90%. Based on the comprehensive optimisation parameters explored, the reaction conditions implemented in **Entry 10** were deemed ideal and used in all subsequent monoalkylation reactions.

After we had ascertained the optimal reaction conditions described above, compound **10** was accordingly monoalkylated with the previously synthesized tosylated glycol ethers **13** – **15**, illustrated in **Scheme 4.3**, to afford compounds **16**, **18** and **19** in good yield. Following this, acylation of these monoalkylated aniline fragments was executed seamlessly using acryloyl chloride and K₂CO₃ in THF at low temperature. This afforded the *N*-SEM protected, Michael acceptor containing compounds **20** – **22** in excellent yields and set the stage for deprotection of all previously synthesised compounds. *N*-SEM deprotection was accomplished by stirring in a 3:1 ratio of DCM/trifluoroacetic acid, followed by further stirring in a 1:1 ratio of a 1 M KOH/THF mixture after removal of the trifluoroacetic acid *in vacuo*. These conditions were applied to compound **10** to afford the simple aniline bearing pyrrolopyrimidine scaffold **23**, as well as to **16**, **18**, **19** and **20** – **22**, to afford the final compounds of series A and series B respectively, as seen in **Scheme 4.3**. Owing to the structural similarity of the final compounds of series A and B, and in the interest of saving time for the reader, the ¹H and ¹³C spectra of compound **27** alone will be depicted and scrutinized. This system will be followed for the rest of this chapter and for the remaining series C - H in our combinatorial library. Full characterization of all compounds not discussed in this chapter, including ¹H, ¹³C and HRMS analysis, can be accessed in the supplementary information (**Section 4.10**).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



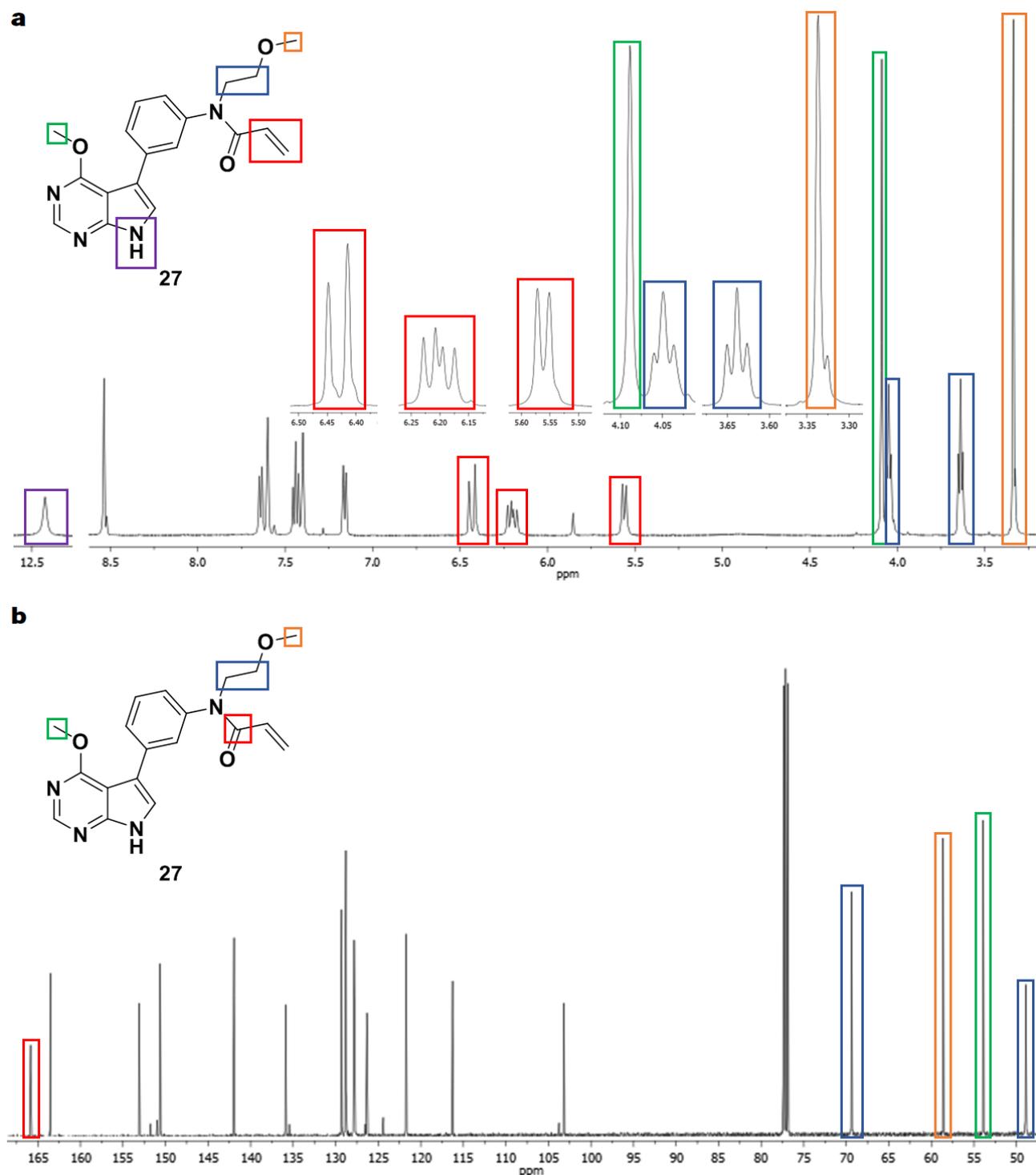
Scheme 4.3: Final synthetic steps including monoalkylation, acylation and deprotection to afford compounds **24** – **29** of series A and B. Reagents and conditions: i) **13** – **15** (0.66 equiv.), K_2CO_3 (1.5 equiv.), DMF, 80 °C, 32 h; ii) acryloyl chloride (1.2 equiv.), K_2CO_3 (1.5 equiv.), THF, –10 °C – rt, 6 h; iii) DCM/TFA (3:1), rt, 6 h, then 1 M KOH/THF (1:1), rt, 4 h.

Examination of the 1H NMR spectrum, shown in **Figure 4.7a**, indicated formation of final compound **27**. In the upfield region, we saw two singlets at δ 3.34 and 4.09 ppm, corresponding to the glycol (orange) and pyrimidine (green) methoxy groups respectively. In between these two sets of singlets, we saw two triplets at δ 3.64 and 4.05 ppm, which correlated to the ethylene bridge of the glycol chain (dark blue). Moving downfield, the characteristic acrylamide signals were seen between δ 5.55 and 6.41, comprising of two doublets and a doublet of doublets (red). Owing to the absence of any silyl group signals at δ 0.00 ppm, and the presence of a broad singlet at δ 12.4 ppm (purple) which belonged to the pyrrole amine proton, we could consider the deprotection of the *N*-SEM protecting group as going to completion.

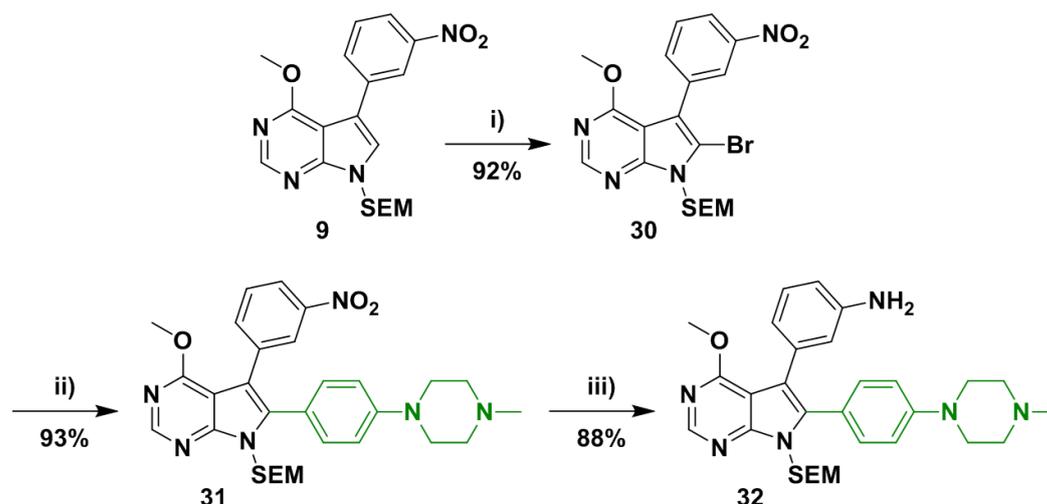
Inspection of the ^{13}C NMR spectrum in **Figure 4.7b** further reinforced the structural findings of the 1H spectrum, with four distinct peaks present in the aliphatic region. Working from NMR spectra of precursory structures and visual comparison of the peak intensity, the pyrimidine (**Figure 4.7b**, green) and glycol methoxy carbons signals could be assigned to δ 53.9 and 58.6 ppm respectively. The glycol bridge carbon signals (dark blue) were found at δ 48.9 and 69.4 ppm. The appearance of a signal in the furthest downfield region at δ 165.9 ppm signified the carbonyl carbon of the acrylamide.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

All undiscussed proton signals were correctly accounted for, yielding anticipated integration values and splitting patterns. Furthermore, 18 of the 19 carbon signals could be accounted for, while the remaining absent signal was ascribed to a quaternary carbon and the slow relaxation times of these centres. Additionally, use of HRMS analysis presented a $[M+H]^+$ ion of 353.1612, with a calculated mass of 353.1614, providing final confirmation for the successful synthesis of compound **27**. Similar results were obtained for the remaining final compounds of series A and B, affirming the successful synthesis thereof.



4.5.3 Synthesis of Series C and D

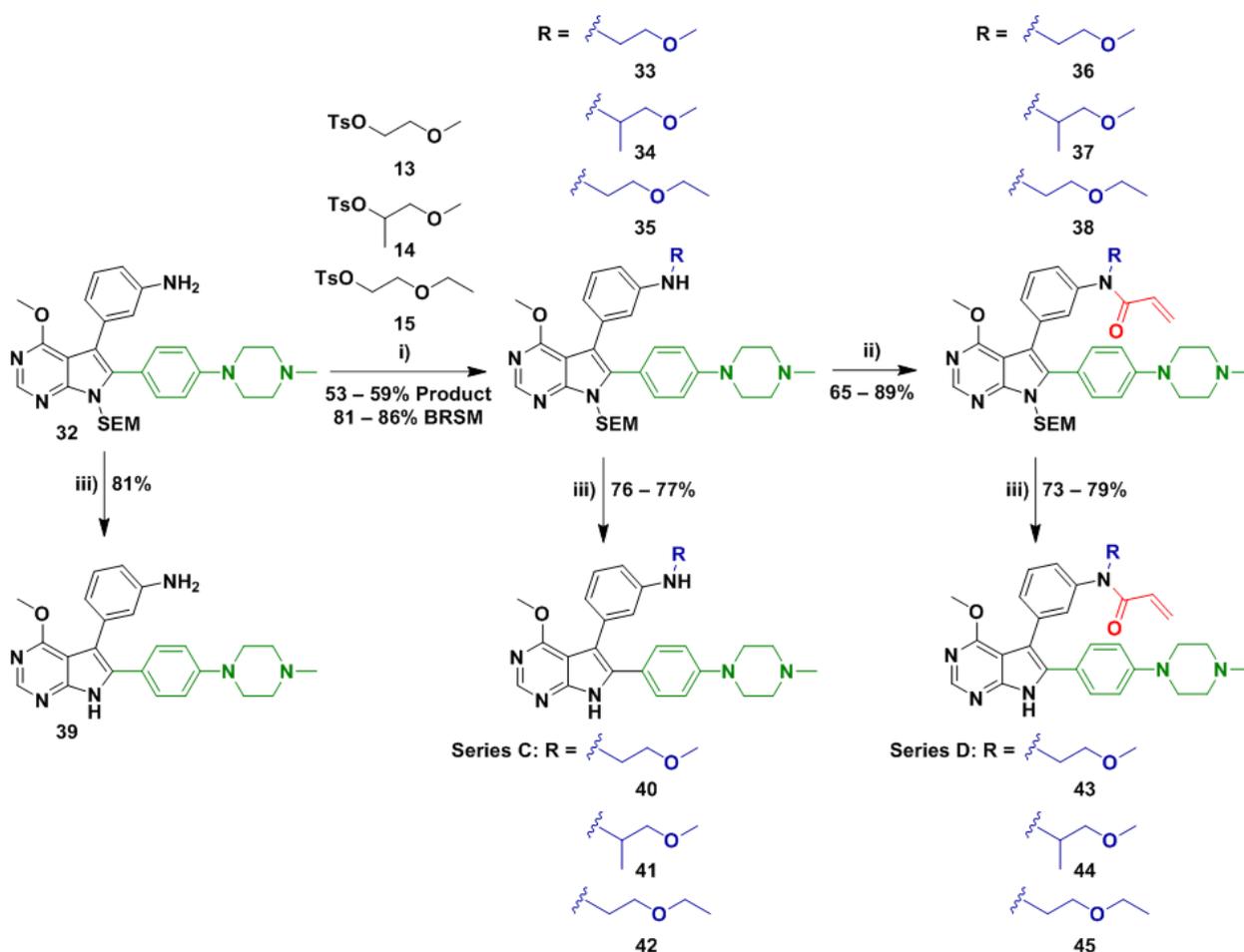


Scheme 4.4: Three step synthesis of the solubilising group containing pyrrolopyrimidine core structure **32**. Reagents and conditions: i) NBS (1.05 equiv.), DCM, rt, 16 h; ii) 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine (1.2 equiv.), Pd(PPh₃)₄ (5 mol%), Na₂CO₃ (2.5 equiv.), dioxane/H₂O (3:1), 80 °C, 16 h; iii) Fe (5 equiv.), NH₄Cl (1 equiv.), EtOH/H₂O (1:1), ultrasonication at 60 °C, 6 h.

Following the successful synthesis of series A and B, carried out with the established monoalkylation reaction conditions, we set about synthesising series C and D which would contain the phenyl methyl-piperazine solubilising group shown in green in **Scheme 4.4** above. Starting from the previously synthesised compound **9**, selective bromination at the 6-position was accomplished using NBS as our source of electrophilic bromine in DCM to afford compound **30** in excellent yield. This allowed for coupling of compound **30** with 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine using the optimized Suzuki-Miyaura reaction conditions, furnishing compound **31** in outstanding yield with the solubilising group shown in green. Lastly, reduction of the aromatic nitro group was achieved using iron and NH₄Cl under sonication at 60 °C to supply us with the core structure **32**, which would be used for further derivatization.

The concluding steps in the synthesis of the final compounds of series C and D were undertaken in an analogous manner to that of series A and B and can be seen in **Scheme 4.5**. These steps included the monoalkylation of core structure **32** with the tosylated glycol ethers **13** – **15**, to afford the intermediate compounds **33** – **35**. This was followed by the acylation of these monoalkylated aniline fragments with acryloyl chloride to install the Michael acceptor moiety and provide the *N*-SEM protected final compounds **36** – **38**. Deprotection of the *N*-SEM functionality was accomplished using TFA and KOH successively for all compounds, to provide the final compounds **40** – **42** and **43** – **45** of series C and D respectively. ¹H and ¹³C NMR spectroscopic analysis of final compound **43** follows.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



Scheme 4.5: Final synthetic steps including monoalkylation, acylation and deprotection to afford final compounds **40** – **45** of series C and D. Reagents and conditions: i) **13** – **15** (0.66 equiv.), K_2CO_3 (1.5 equiv.), DMF, $80\text{ }^\circ\text{C}$, 32 h; ii) acryloyl chloride (1.2 equiv.), K_2CO_3 (1.5 equiv.), THF, $-10\text{ }^\circ\text{C}$ – rt, 6 h; iii) DCM/TFA (3:1), rt, 6 h, then 1 M KOH/THF (1:1), rt, 4 h.

Inspection of the ^1H and ^{13}C NMR spectra of compound **43**, which can be seen in **Figure 4.8** on the following page, revealed related signals corresponding to the same structural elements present in final compound **27**, which appeared at similar chemical shifts for compound **43**. For the ^1H NMR spectrum, shown in **Figure 4.7a**, this included: the two singlets of the glycol (orange) and pyrimidine (green) methoxy groups, two sets of triplets which correlated to the ethylene groups of the glycol chain (dark blue), the three unsaturated protons of the Michael acceptor visible as three doublet of doublets (red) and the downfield pyrrole amine proton (purple). Furthest upfield at δ 2.42 ppm we saw the appearance of a tall singlet assigned to the *N*-methyl group of the piperazine group (yellow). Furthermore, the emergence of two triplets integrating for 4H at δ 2.46 and 3.28 ppm made up the remainder of the cyclic piperazine (light blue).

Similarly, the previously assigned carbon signals which were discussed for compound **27** were evident in the ^{13}C NMR spectrum of compound **43**, shown in **Figure 4.8b**. These included the two methoxy methyl carbons (green and orange), glycol ethylene carbons (dark blue) and the carbonyl carbon of the acrylamide (red). New peaks included the carbon at δ 46.0 ppm, which was assigned to the *N*-methyl piperazine carbon (yellow), and the four remaining carbons of the cyclic piperazine structure, which appeared as a pair of high-intensity signals at δ 48.0 and 54.8 ppm owing to the symmetry of the solubilising group (light blue).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

This was equally apparent for the phenyl linker, whose two signals occurred at δ 115.6 and 129.0 ppm in the aromatic region (purple). As with analysis of compound **27**, all unassigned protons and carbons within the spectra were correctly accounted for, with use of HRMS analysis presenting a $[M+H]^+$ ion of 527.2773, with a calculated mass of 527.2777. This providing final confirmation for the successful synthesis of compound **43**, and in a similar manner, the other compounds comprising series C and D.

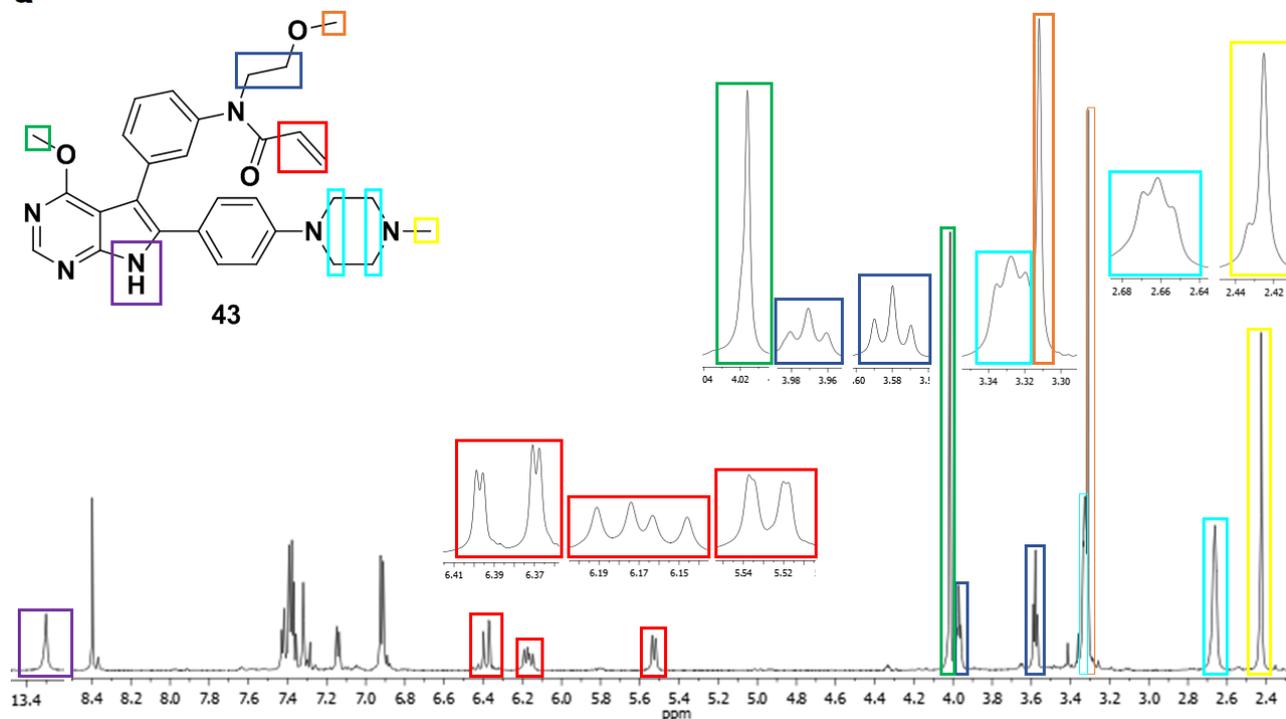
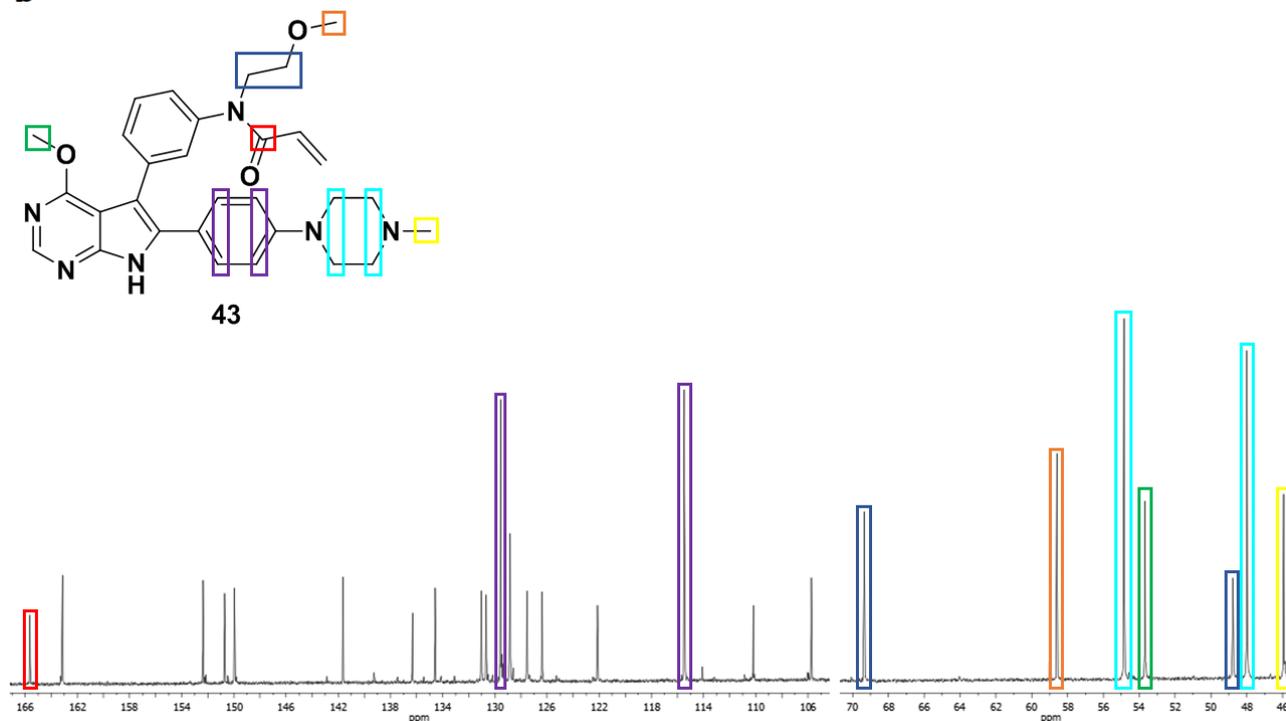
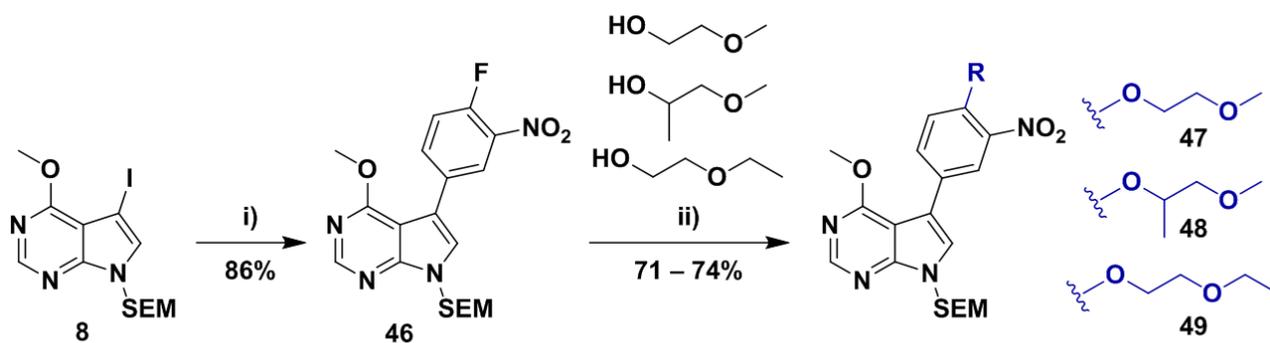
a**b**

Figure 4.8: a) ^1H and b) ^{13}C NMR spectra of final compound **43**.

4.5.4 Synthesis of Series E and F



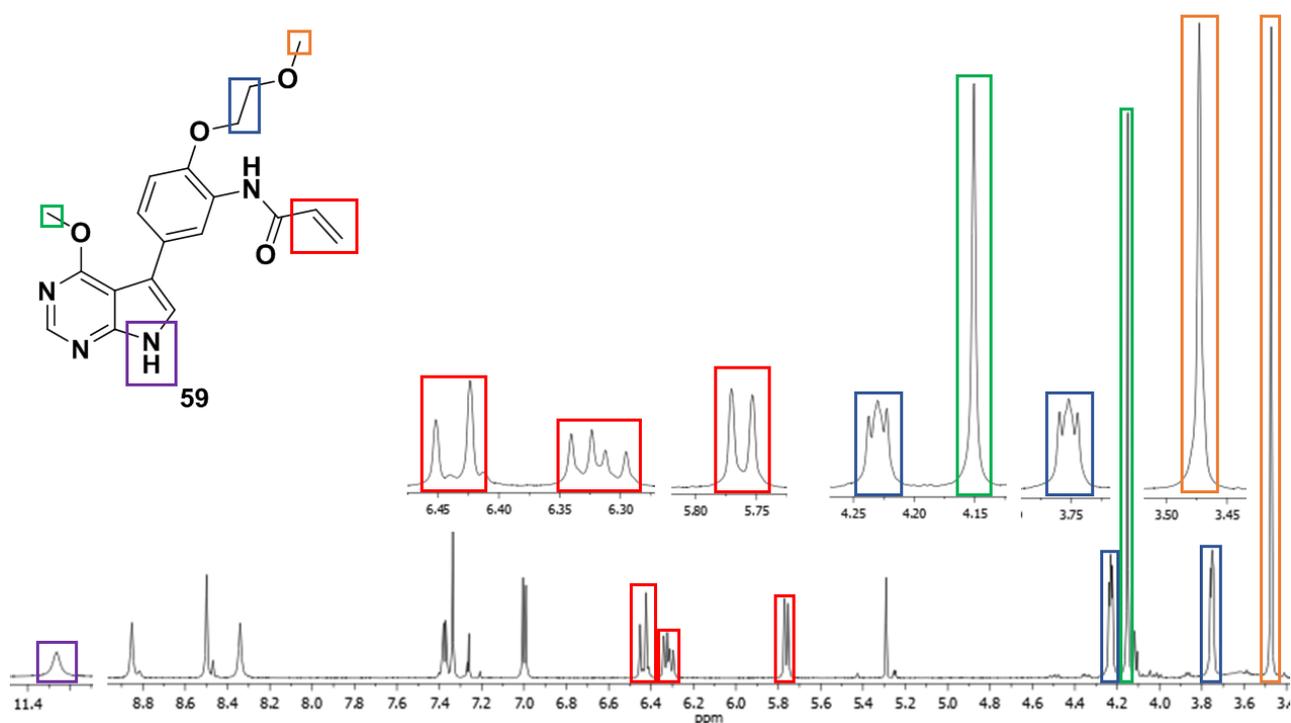
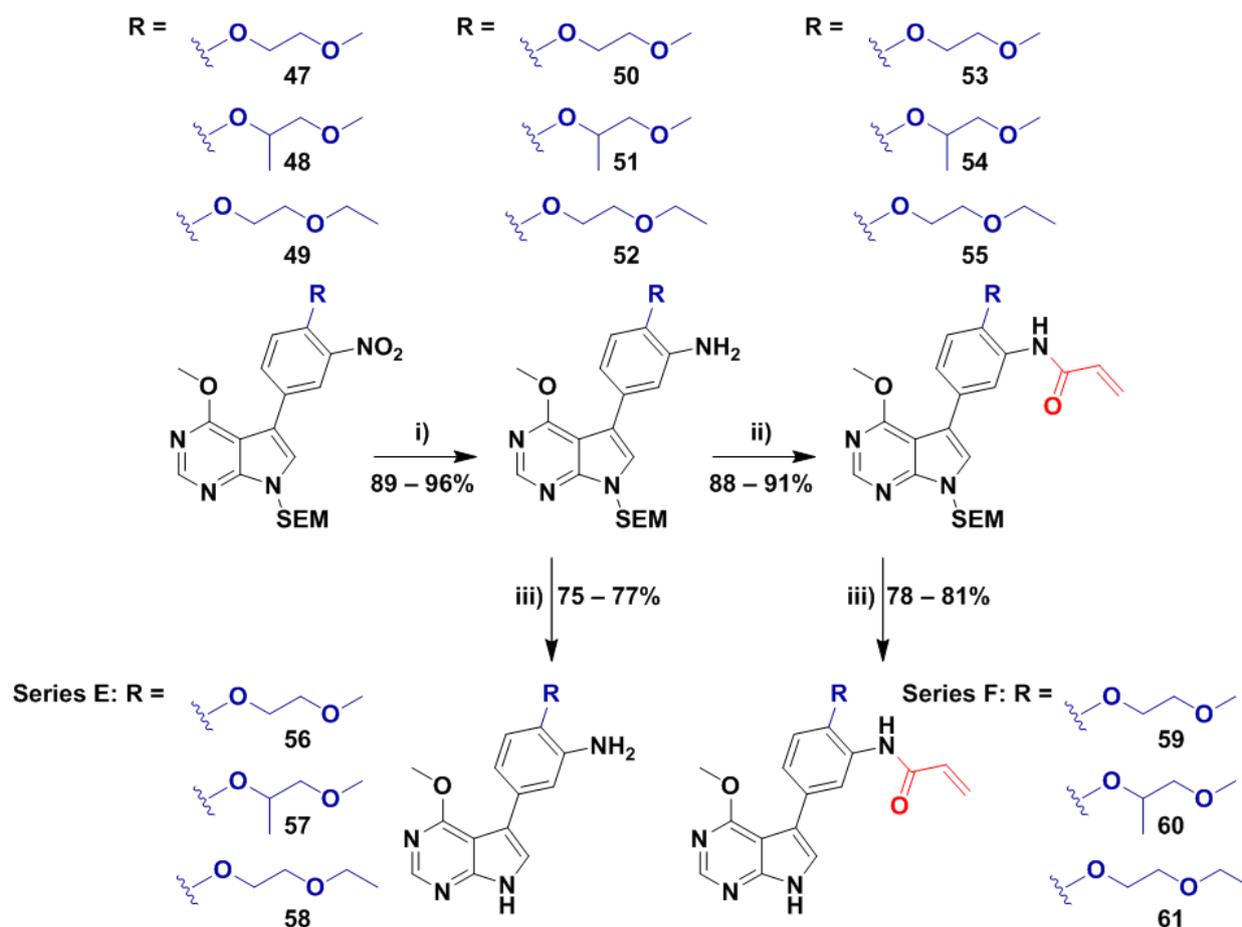
Scheme 4.6: Two step synthesis of the glycol ether-containing pyrrolopyrimidine intermediates **47** – **49**. Reagents and conditions: i) (4-fluoro-3-nitrophenyl)boronic acid (1.2 equiv.), Pd(PPh₃)₄ (5 mol%), Na₂CO₃ (2.5 equiv.), dioxane/H₂O (3:1), 80 °C, 16 h; iii) glycol ether derivative (1.2 equiv), NaH (1.7 equiv.), THF, –10 °C – rt, 16 h.

The final compounds comprising series E – H required installation of the glycol chain in an *ortho*-position to the aniline or Michael acceptor moiety. To accomplish this, we envisaged making use of a nucleophilic aromatic substitution reaction. By coupling the commercially available (4-fluoro-3-nitrophenyl)boronic acid to the previously synthesised **8**, we would place a fluorine atom in the correct position to undergo the S_NAr reaction and allow for introduction of the glycol ether chains, as outlined in **Scheme 4.6** above. As fluorine is highly electronegative, the reaction rate for nucleophilic attack would be increased by its activation of the electrophilic aromatic carbon. Furthermore, this reaction centre would be *ortho* to the strongly electron-withdrawing nitro group, further activating the ring towards nucleophilic attack by the alkyl alcohol. Carrying out the coupling of compound **8** to (4-fluoro-3-nitrophenyl)boronic acid under the optimised Suzuki-Miyaura conditions provided compound **46** in high yield. Interestingly, the ¹³C NMR spectrum of **46** showed splitting for the aromatic carbons neighbouring the fluorine atom, with a total of five signals split into doublets. This also indicated that the fluorine atom had survived the high temperature and palladium catalyst used in the prior boronic acid coupling. Subsequent S_NAr reaction of the commercially available alkyl alcohols shown in **Scheme 4.6** with **46**, using sodium hydride as base, furnished the intermediate compounds **47** – **49** in good yield.

With the above synthesized intermediates in hand, we embarked on the final steps in the synthesis of series E and F, shown in **Scheme 4.7** on the next page. Aromatic nitro reduction of compounds **47** – **49**, using the previously established conditions of iron and ammonium chloride, provided the *ortho*-glycol aniline fragments **50** – **52** in excellent yield. Subsequent acylation of these aniline fragments with acryloyl chloride installed the Michael acceptor moiety and afforded the protected final compounds **53** – **55**. Deprotection of the *N*-SEM group of both sets of compounds, **50** – **52** and **53** – **55**, furnished the final compounds **56** – **58** and **59** – **61**, of series E and F, respectively.

Due to the structural similarity of these compounds to those synthesised in series A and B, only the ¹H NMR spectrum of compound **59** will be interpreted and displayed on the following page in **Figure 4.9**. Comparable to the spectral analysis of compound **27**, compound **59** retained the two signals of the glycol and 4-pyrimidine position methoxy protons (orange and green respectively), with the three sets of signals between δ 5.75 and 6.45 ppm corresponding to the acryl group protons (red) and the pyrrole amine proton far downfield at δ 11.3 ppm.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



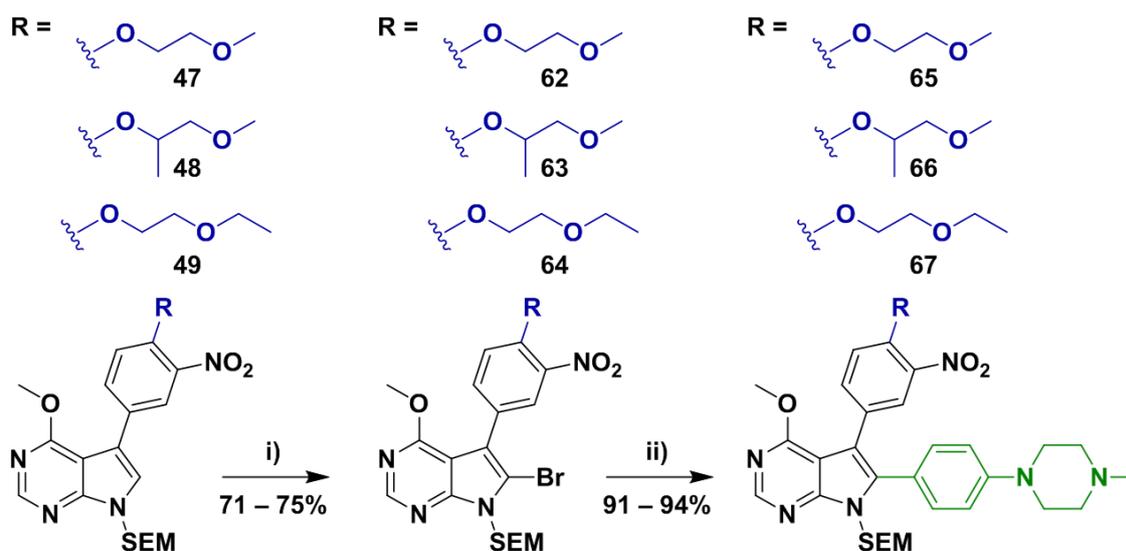
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

However, where the glycol ethylene proton peaks occurred between the two methoxy signals in compound **27**, we saw a shift in one ethylene triplet (dark blue) to a position further downfield than the pyrimidine methoxy (green) in the spectrum of compound **59**. This could be attributed to the additional electron-withdrawing effects of direct substitution on the aromatic ring and the presence of the extra electronegative oxygen atom in the glycol ether chain. All other aromatic protons and carbon signals within the ^{13}C NMR spectrum were fully accounted for. Finally, use of HRMS analysis presented a $[\text{M}+\text{H}]^+$ ion of 369.1560, with a calculated mass of 369.1563, providing confirmation for the successful synthesis of compound **59** and, in a similar fashion, the remaining compounds of series E and F.

4.5.5 Synthesis of Series G and H

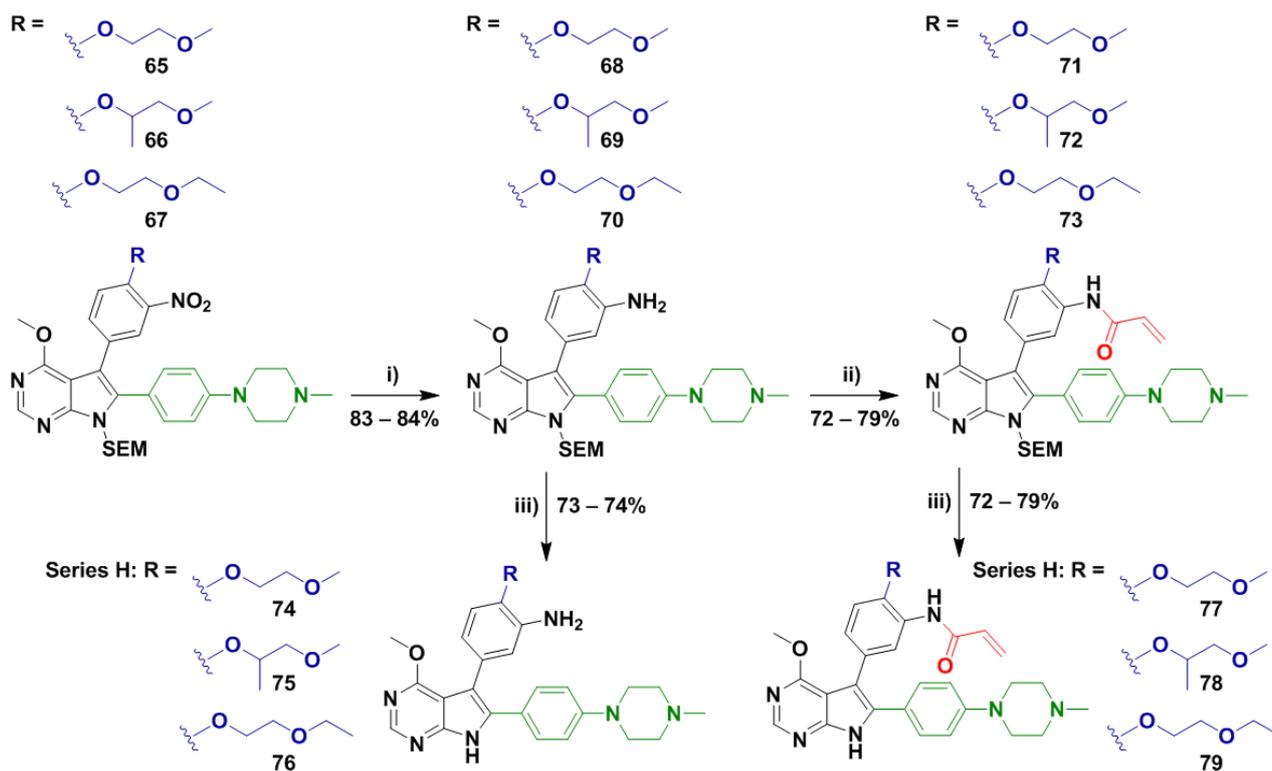
With the *ortho*-glycol derivatised, aromatic nitro-containing intermediates **47** – **49** in hand, the first steps in the synthesis of series G and H was incorporation of the phenyl methyl-piperazine solubilising group onto these scaffolds, as illustrated in **Scheme 4.8**. This required selective bromination at the 6-position, which was accomplished using NBS in DCM, affording compounds **62** – **64** in moderately good yield. Next, the Suzuki-Miyaura reaction was successfully undertaken, coupling the pinacol boronic ester of the solubilising group to compounds **62** – **64**, thereby yielding the intermediate compound **65** – **67** shown below in **Scheme 4.8**.

The concluding steps in the synthesis of the ultimate two series of compounds followed a similar pattern used previously and is shown in **Scheme 4.9** on the following page. This included reduction of the aromatic nitro group, subsequent acylation of the formed aniline and deprotection. Thus, the reduction of compounds **65** – **67** was achieved by sonication with a mixture of iron and ammonium chloride, furnishing the three aniline fragments **68** – **70**, followed by acylation with acryloyl chloride to provide the protected compounds **71** – **73**. Lastly, deprotection of the *N*-SEM functionalised compounds **68** – **73**, using TFA and KOH, supplied us with the final compounds **74** – **76** and **77** – **79**, comprising of series G and H respectively.



Scheme 4.8: Two step synthesis of the solubilising group containing pyrrolopyrimidine intermediates **65** – **67**. Reagents and conditions: i) NBS (1.05 equiv.), DCM, rt, 16 h; ii) 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine (1.2 equiv.), $\text{Pd}(\text{PPh}_3)_4$ (5 mol%), Na_2CO_3 (2.5 equiv.), dioxane/ H_2O (3:1), 80 °C, 16 h;

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



Scheme 4.9: Final synthetic steps including reduction, acylation and deprotection to afford compounds 74 – 79 of series G and H. Reagents and conditions: i) Fe (5 equiv.), NH₄Cl (1 equiv.), EtOH/H₂O (1:1), ultrasonication at 60 °C, 6 h; ii) acryloyl chloride (1.2 equiv.), K₂CO₃ (1.5 equiv.), THF, –10 °C – rt, 6 h; iii) DCM/TFA (3:1), rt, 6 h then 1 M KOH/THF (1:1), rt, 4 h.

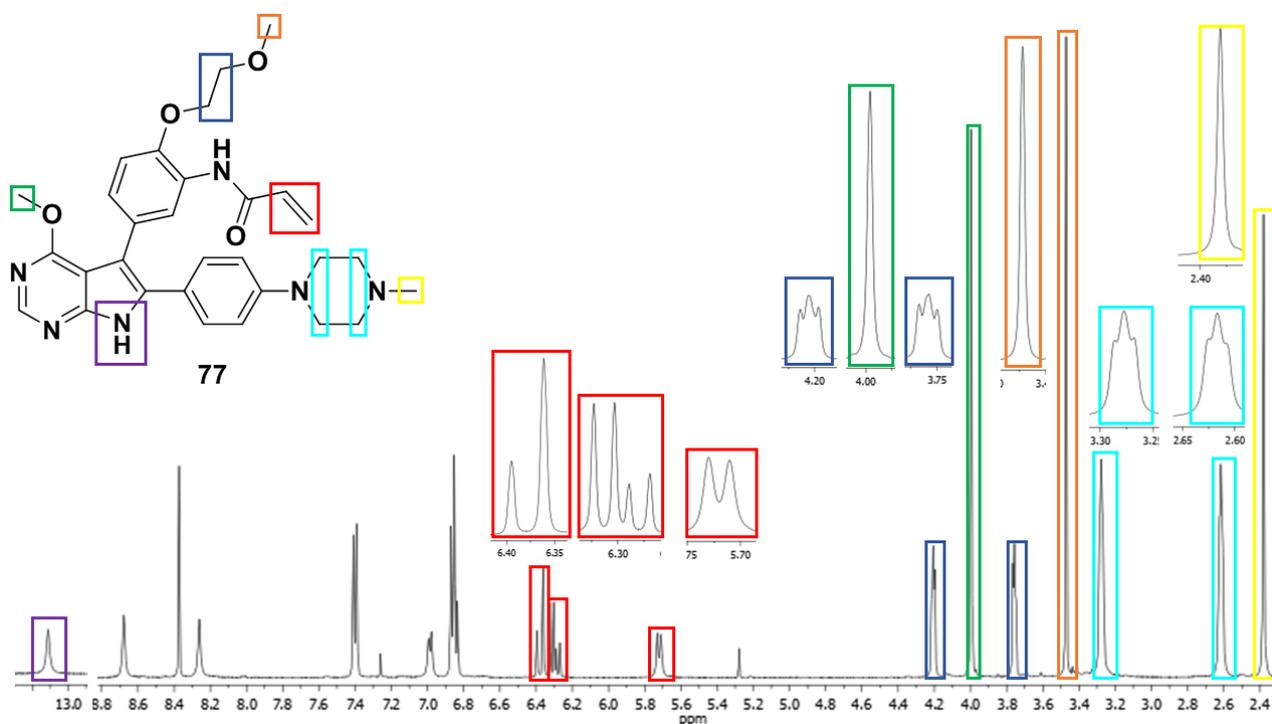


Figure 4.10: ¹H NMR spectra of final compound 77.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

As implemented before with compound **59** of series F, the ^1H NMR spectrum of final compound **77** will be discussed exclusively and is displayed in **Figure 4.10** on the previous page. The shared structural signatures found in all compounds analysed thus far were evident in the ^1H spectrum of compound **77**. Furthest upfield we saw the *N*-methyl piperazine singlet (yellow) and the two triplets comprising the remainder of the piperazine heterocycle (light blue). Moving downfield, the glycol methoxy peak (orange) occurred at δ 3.45 ppm and, following the substitution pattern of that seen in compound **59**, we saw the two triplets of the ethylene glycol bridge (dark blue) flanking the pyrimidine methoxy (green). The acryl group protons (red) could be seen in the region between δ 5.70 and δ 6.40 ppm and the pyrrole amine proton appeared as a broad singlet at δ 13.1 ppm. In line with previous spectroscopic analyses, all undiscussed aromatic protons and carbon signals within the ^{13}C NMR spectrum were fully accounted for. Furthermore, use of HRMS analysis presented a $[\text{M}+\text{H}]^+$ ion of 543.2717, with a calculated mass of 543.2720, providing final confirmation for the successful synthesis of compound **77**. These analytical considerations were reflected in the remaining compounds of series G and H, affirming the successful synthesis thereof.

This concluded the synthesis of our combinatorial target library, affording 26 final compounds which were evaluated in biochemical and cellular assays and will be discussed comprehensively in the next section.

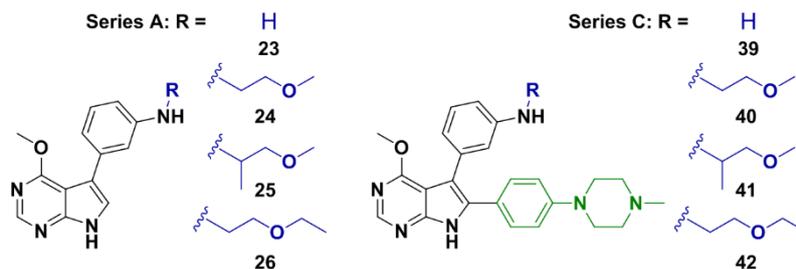
4.6 Biochemical and Cellular Evaluation

Owing to the multitude of compounds encompassing the synthesised library, we chose to analyse the biochemical and cellular data acquired by firstly discussing and comparing similar structural features in a series vs series manner. This would be followed by selecting the most efficacious inhibitors in each of these comparative analyses and combining them to provide an overview of the results. By this method, we first compared compounds which do not contain the electrophilic acrylamide functionality, to determine the factors that promote favourable reversible inhibition. This means that series A vs C (directly substituted, without vs with solubilising group) and series E vs G (*ortho*-substitution, without vs with solubilising group) were scrutinized. Based on these results, the most potent reversible inhibitors, and their desirable structural features, could be identified. This same process was repeated for the series of compounds which contain the Michael acceptor moiety and could potentially act in an irreversible manner. Lastly, we provided an overview of the most efficacious reversible vs irreversible inhibitors, placing emphasis on the activity of the compounds against the double and triple mutant variant.

To determine efficacy in a biochemical setting, the compounds were evaluated against the wild-type, activating L858R, double L858R/T790M and triple L858R/T790M/C797S mutant variants of the enzyme. For the cellular tests, we employed a wild-type-bearing cell line A431, the EGFR-L858R/T790M drug-resistant cell line H1975 and clinically relevant osimertinib resistant cell line EGFR-PC9/T790M/C797S. Compound assessment was carried out either by Dr. Jonas Lategahn or Dr. Marina Keul, with all measurements being executed in at least triplicate. The respective IC_{50} and EC_{50} values of osimertinib and lead compound **1** accompany the obtained results in each report.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

4.6.1 Comparison of Reversible Inhibitors

Table 4.2: Comparison of the scaffolds **23** and **39** and inhibitors **24 – 26** and **40 – 42** from series A and C respectively.

Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M	L858R/ T790M/ C797S	A431	H1975	PC9/ T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	113 ± 19	756 ± 340	16.0 ± 5.00	3109 ± 648
1	0.41 ± 0.01	0.12 ± 0.01	0.24 ± 0.02	50 ± 19	638 ± 247	60.0 ± 8.00	-
23	1160 ± 132	254 ± 50.3	-	14457 ± 9599	25330 ± 3110	19630 ± 4120	30 ± 0
24	2990 ± 231	639 ± 316	720 ± 131	7825 ± 404	30000 ± 0	30000 ± 0	-
25	2785 ± 424	405 ± 67.1	726 ± 23.1	10803 ± 845	30000 ± 0	30000 ± 0	-
26	1854 ± 238	369 ± 55.3	569 ± 72.0	4451 ± 101	30000 ± 0	30000 ± 0	-
39	2.33 ± 0.41	0.616 ± 0.06	-	12.04 ± 9.72	1209 ± 224	1385 ± 452	1603 ± 199
40	3.63 ± 1.54	0.55 ± 0.23	1.75 ± 0.98	1.50 ± 0.08	1128 ± 150	1471 ± 130	2489 ± 490
41	3.42 ± 2.01	0.69 ± 0.40	1.70 ± 1.05	9.27 ± 6.31	4332 ± 274	2925 ± 730	4135 ± 480
42	6.20 ± 2.72	0.71 ± 0.29	2.26 ± 1.41	10.4 ± 10.9	2783 ± 1270	3416 ± 1210	-

Examination of the activity profiles for the directly substituted glycol compounds of series A and C, as well as their unsubstituted aniline counterparts **23** and **39**, revealed some interesting trends with regards to the presence or absence of the solubilising group and is shown in **Table 3.2** above. The unsubstituted pyrrolopyrimidine aniline compound **23**, which does not bear the solubilising group, showed very little reversible inhibition both in a biochemical and cellular setting. This trend was followed in the glycol substituted compounds of series A (**24 – 26**), which also do not contain the solubilising group, showing high micromolar biochemical activity against the triple mutant and displaying none against the double mutant in a cellular evaluation. This warranted the absence of screening this series against the triple mutant cell line. In contrast, we saw a large spike in activity for the compounds which do contain the solubilising group. The unsubstituted aniline pyrrolopyrimidine compound **39** displayed a dramatic increase in activity against all variants biochemically when compared with its counterpart **23**. An IC₅₀ value of 12 nM was found against the triple mutant - a 10- and 4-fold superiority over osimertinib and the lead compound **1** respectively. This trend was followed in the cellular context, with low micromolar EC₅₀ values in all cell lines and a 2-fold increase over osimertinib against the triple mutant. This highlighted the exemplary reversible binding affinity of the core pyrrolopyrimidine scaffold and the importance of the solubilising group in both cell permeability and promoting potent reversible interactions by anchoring the molecule in the correct orientation within the active site.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

The weak activity against the EGFR-L858R-T790M mutant was largely expected, owing to the absence of the electrophilic acrylamide and the inability to covalently inhibit the enzyme.

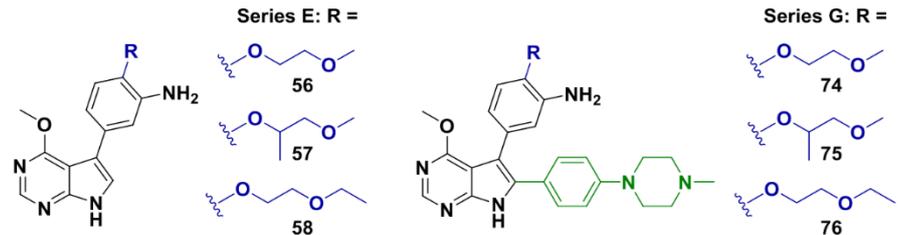
The directly substituted glycol variants of compound **39**, which make up series C, showed similarly impressive reversible activity. Biochemically, compounds **40** – **42** showed sub-nanomolar IC₅₀ values for the activating L858R mutant and single digit IC₅₀ values for the T790M double mutant, even with the exclusion of the Michael acceptor. These compounds retained their remarkable activity against the triple mutant, with compound **40** and **41** showing single digit IC₅₀ values of 1.50 and 9.27 nM respectively. We discerned that increased length and substitution of the glycol chain led to a loss of activity, as was exemplified by the decreased activity for compounds **41** and **42** when compared with **40**. This notion was further supported in the cellular evaluation of compounds **40** and **41** against the triple mutant, displaying a similar drop in activity with increasing substitution. Importantly, both compounds **39** and **40** showed improved cellular activity over osimertinib against this mutant variant.

When comparing the biochemical results of the core compound **39** with the 2-methoxyethyl substituted pyrrolopyrimidine **40**, we saw that compound **40** had a 6-fold greater activity against the triple mutant. Conversely, the same comparison in a cellular evaluation revealed compound **39** showing slightly better activity than its substituted counterpart **40**. This alluded to the potential instability and ill tolerance of the glycol chain moieties within the cell. From the above comparison we could thus conclude that the solubilising group is extremely important for reversible inhibition activity, as well as shorter, less-substituted glycol chains being favoured.

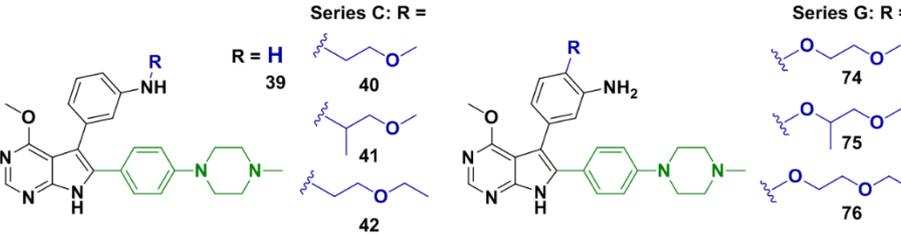
Similarly, we wished to see the effect of *ortho*-substitution of the glycol chain moieties on our pyrrolopyrimidine based inhibitors activity. The results, which can be seen on the following page in **Table 4.3**, mirror the findings of **Table 4.2**, with a large discrepancy in inhibition based on the absence or presence of the solubilising group, poor activity against the T790M double mutant and strong biochemical inhibition for series G against the C797S triple mutant. Furthermore, the trend of decreasing activity with increased glycol chain length and substitution was repeated, with compounds **75** and **76** (additional methyl substituent and longer glycol chain respectively) showing a decrease in biochemical efficacy across all variants of the enzyme. This reinforced the conclusions drawn in our previous comparison. Owing to its biochemical efficiency, we chose to evaluate compound **74** against the triple mutant cell line which displayed an EC₅₀ value of roughly 3 μM and a minor increase in activity over osimertinib.

Lastly, a comparison was drawn between the most efficacious directly and *ortho*-substituted reversible inhibitors to determine the optimal position of the glycol chain. These included the core scaffold **39** and compounds **40** – **42** and **74** – **76** of series C and G respectively. Visual inspection of these values in **Table 4.4** on the following page, showed that the directly substituted series C exhibited a more favourable activity profile than series G, with compound **40** foremost of all the reversible inhibitors synthesised. Between these series, compound **40** exhibited the greatest inhibitory activity against the triple mutant both biochemically and in a cellular context and was the only molecule to show selectivity for mutant variants of EGFR over the wild-type enzyme. However, it is important to note that the unsubstituted scaffold **39** retained the highest activity against the triple mutant cell line, which unfortunately underlines the inefficiency of our synthesised glycol-chain derivatised pyrrolopyrimidines in a cellular setting. With the reversible inhibitor's efficacy investigated, we turned our attention towards the synthesised Michael-acceptor containing compounds.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Table 4.3: Comparison of reversible inhibitors **56 – 58** and **74 – 76** from series E and G respectively.


Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M	L858R/ T790M/ C797S	A431	H1975	PC9/ T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	113 ± 19	756 ± 340	16.0 ± 5.00	3109 ± 648
1	0.41 ± 0.01	0.12 ± 0.01	0.24 ± 0.02	50 ± 19	638 ± 247	60.0 ± 8.00	-
56	4762 ± 478	335 ± 190	499 ± 139	20000 ± 0	30000 ± 0	30000 ± 0	-
57	336 ± 106	309 ± 171	1203 ± 357	20000 ± 0	30000 ± 0	30000 ± 0	-
58	996 ± 395	147 ± 97	305 ± 192	20000 ± 0	13040 ± 3720	14200 ± 3820	-
74	2.23 ± 1.70	0.50 ± 0.38	1.45 ± 1.31	10.1 ± 7.51	2932 ± 1930	3833 ± 530	3068 ± 450
75	4.65 ± 1.90	0.78 ± 0.68	2.78 ± 1.59	25.6 ± 14.2	3648 ± 190	3233 ± 153	-
76	4.21 ± 1.88	0.86 ± 0.38	2.48 ± 1.39	36.3 ± 21.1	1264 ± 80	1936 ± 720	-

Table 4.4: Comparison of reversible inhibitors **40 – 42** and **74 – 76** from series E and G respectively.


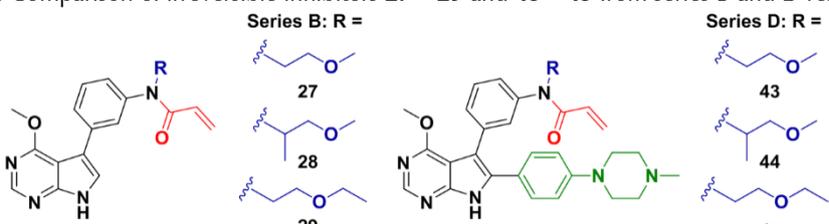
Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M	L858R/ T790M/ C797S	A431	H1975	PC9/ T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	113 ± 19	756 ± 340	16.0 ± 5.00	3109 ± 648
1	0.41 ± 0.01	0.12 ± 0.01	0.24 ± 0.02	50 ± 19	638 ± 247	60.0 ± 8.00	-
39	2.33 ± 0.41	0.616 ± 0.06	-	12.04 ± 9.72	1209 ± 224	1385 ± 452	1603 ± 199
40	3.63 ± 1.54	0.55 ± 0.23	1.75 ± 0.98	1.50 ± 0.08	1128 ± 150	1471 ± 130	2489 ± 490
41	3.42 ± 2.01	0.69 ± 0.40	1.70 ± 1.05	9.27 ± 6.31	4332 ± 274	2925 ± 730	4135 ± 480
42	6.20 ± 2.72	0.71 ± 0.29	2.26 ± 1.41	10.4 ± 10.9	2783 ± 1270	3416 ± 1210	-
74	2.23 ± 1.70	0.50 ± 0.38	1.45 ± 1.31	10.1 ± 7.51	2932 ± 1930	3833 ± 530	3068 ± 450
75	4.65 ± 1.90	0.78 ± 0.68	2.78 ± 1.59	25.6 ± 14.2	3648 ± 190	3233 ± 153	-
76	4.21 ± 1.88	0.86 ± 0.38	2.48 ± 1.39	36.3 ± 21.1	1264 ± 80	1936 ± 720	-

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

4.6.2 Comparison of Potentially Irreversible Inhibitors

The same method of comparison used previously for the reversible inhibitors was applied to the potentially irreversible compounds. Accordingly, the activities of the directly substituted glycol-moiety, acrylamide-containing compounds, with or without the solubilising group, can be seen in **Table 4.5** below. The first observable trend concerned the effects of the solubilising group in series B and D. Once again, we saw the distinct requirement of the solubilising group in sustaining potent inhibitory activity, with compounds **27** – **29** showing poor to almost no activity against EGFR-L858R/T790M biochemically and in cells. We therefore disregarded further investigation and discussion for the compounds of series B. Examination of the inhibitory concentrations for compounds **43** – **45**, which comprise series D, reflected an immense increase in inhibition of the double mutant, with IC₅₀ values in the low picomolar range. This elevated activity suggested irreversible inhibition through the Michael acceptor, with supplementary reversible interactions and binding affinity promoted by the glycol chains. Interestingly, the biochemical results for the activating and double mutant of the irreversible compounds did not follow the trend of decreasing activity with increasing substitution. Compound **44**, which bears the 1-methoxypropan-2-yl chain, displayed the highest activity, with an IC₅₀ of 70 pM, a 3- and 8-fold increase in activity over the lead compound **1** and osimertinib respectively, as well as enhanced selectively over the wild-type enzyme. We believe the unique methyl substituent of this glycol chain could interact sterically with the proximal acryl amide electrophile, promoting a favourable orientation specifically for covalent bond formation with Cys797. In contrast, we saw a reversal of this trend against the triple mutant, with the short chain derivatised compound **43** being the most efficacious inhibitor with an IC₅₀ value of 86 nM. This further substantiated our proposal, as covalent bond formation with this electrophile is not possible with the serine bearing triple mutant, negating the positive effect of the methyl substituent.

Table 4.5: Comparison of irreversible inhibitors **27** – **29** and **43** – **45** from series B and D respectively.



Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M	L858R/ T790M/ C797S	A431	H1975	PC9/ T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	113 ± 19	756 ± 340	16.0 ± 5.00	3109 ± 648
1	0.41 ± 0.01	0.12 ± 0.01	0.24 ± 0.02	50 ± 19	638 ± 247	60.0 ± 8.00	-
27	357 ± 98.9	103 ± 12.8	297 ± 102	20000 ± 0	25540 ± 6820	23790 ± 4510	-
28	774 ± 53.1	199 ± 24.7	774 ± 149	16483 ± 989	21050 ± 7910	27020 ± 3370	-
29	443 ± 50.8	127 ± 13.5	379 ± 126	20000 ± 0	27930 ± 3800	24680 ± 3670	-
43	0.41 ± 0.02	0.11 ± 0.07	0.12 ± 0.04	86.1 ± 55.8	579 ± 160	120 ± 60	-
44	0.41 ± 0.01	0.05 ± 0.04	0.07 ± 0.05	185 ± 64.7	1638 ± 670	190 ± 116	17679 ± 2070
45	0.41 ± 0.05	0.14 ± 0.06	0.19 ± 0.13	141 ± 72.5	1503 ± 800	277 ± 240	-

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Moving to the cellular evaluation, we witnessed a slight loss of activity for the compounds of series D in the T790M mutation-bearing H1975 cell line, when measured against lead compound **1**. This was a disappointing result, as we anticipated an increase in activity for these compounds, as exhibited by their biochemical data. We surmised that this decline in activity was consequent of the cell's intolerance towards the glycol moiety. Considering its superior cellular selectivity profile and biochemical inhibition properties, we incubated compound **44** within C797S triple mutant bearing cells but were dismayed to see a dramatic drop in activity, resulting in an EC_{50} value of over $17 \mu\text{M}$. While a loss of activity for the resistant triple mutant was in keeping with our expectations, we did not foresee a 6- and 7-fold decrease when compared to osimertinib and the leading reversible inhibitor **40** respectively. As exemplified previously with our reversible compounds, the cells are averse to the inclusion of the glycol moiety. Additionally, the C797S mutation has shown strong resistance profiles to compounds bearing the acrylamide functionality specifically, such as osimertinib and olmutinib. We therefore speculated that this loss of efficacy could be attributed to a combination of the poor pharmacokinetic character of the glycol chain and unfavourable interactions of the acrylamide group with the mutated serine residue.

Proceeding to the *ortho*-substituted, acrylamide-containing compounds of series F and H, shown in **Table 4.6** on the following page, we see similar trends to those found in previous analyses. While the exclusion of the solubilising group resulted in inactivity of series F in a cellular context, the compounds performed surprisingly well in their biochemical evaluation against the L858R/T790M mutant, with compound **60** having an IC_{50} value of 22 nM . Without the anchoring ability of the solubilising group present, we hypothesized that placement of the glycol group in an *ortho*-position provided sufficient reversible interactions to orient the molecule into the correct position for potential covalent modification of the cysteine residue. The compounds of series H, which included the solubilising group, showed remarkable activity in a biochemical setting with IC_{50} values in the low picomolar range against the double mutant. As expected, we see a sharp decrease in activity against the triple mutant for compounds **77 – 79**. With the glycol chain no longer directly substituted, the methyl substituent of **78** was presumed to be too far from the acrylamide electrophile to have any steric effect. This was reflected in the biochemical data, with compound **77** outperforming both **78** and **79** against both the double and triple mutant. This alluded to less substituted, shorter glycol chains being more advantageous for the *ortho*-substitution.

Focussing on the cellular data, we saw an even larger decrease in EC_{50} activity against the L858R/T790M cell line for compounds **77 – 79**, reflecting the results of series D. Furthermore, the pronounced loss of activity against the triple mutant for compounds **78** and **79**, eliciting high micromolar EC_{50} values, indicated the irrelevance of the glycol chain substitution position when targeting this mutant variant. This further augmented our hypothesis of a poor pharmacokinetic profile for these glycol-derivatised molecules, when in combination with the acrylamide group. Notably, compound **77** showed an exceptional selectivity profile for the double mutant over the wild-type enzyme, with EC_{50} values of 298 nM and $27 \mu\text{M}$ respectively, equating to a 100-fold difference in inhibitory concentration. To conclude, we compared the compounds of series D and H to determine the most efficacious irreversible inhibitor, as seen in **Table 4.7** on the following page. Whilst all compounds showed excellent biochemical inhibition, the directly substituted compounds of series D proved to be predominantly more potent than their *ortho*-positioned counterparts across all assays.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Table 4.6: Comparison of irreversible inhibitors **59** – **61** and **77** – **79** from series F and H respectively.

Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M	L858R/ T790M/ C797S	A431	H1975	PC9/ T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	113 ± 19	756 ± 340	16.0 ± 5.00	3109 ± 648
1	0.41 ± 0.01	0.12 ± 0.01	0.24 ± 0.02	50 ± 19	638 ± 247	60.0 ± 8.00	-
59	91.1 ± 15.6	20.3 ± 13.3	26.8 ± 7.32	20000 ± 0	30000 ± 0	20960 ± 9590	-
60	49.1 ± 5.51	17.7 ± 11.9	22.3 ± 4.98	20000 ± 0	30000 ± 0	9556 ± 3460	-
61	190 ± 24.5	30.5 ± 11.1	83.2 ± 18.4	20000 ± 0	30000 ± 0	30000 ± 0	-
77	0.16 ± 0.12	0.08 ± 0.04	0.08 ± 0.03	387 ± 140	27057 ± 5090	298 ± 250	-
78	0.19 ± 0.18	0.07 ± 0.03	0.14 ± 0.06	544 ± 231	15662 ± 1656	931 ± 688	22881 ± 1350
79	0.38 ± 0.10	0.13 ± 0.10	0.14 ± 0.07	983 ± 315	15917 ± 1626	3352 ± 1890	24606 ± 6367

Table 4.7: Comparison of irreversible inhibitors **43** – **45** and **77** – **79** from series D and H respectively.

Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M	L858R/ T790M/ C797S	A431	H1975	PC9/ T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	113 ± 19	756 ± 340	16.0 ± 5.00	3109 ± 648
1	0.41 ± 0.01	0.12 ± 0.01	0.24 ± 0.02	50 ± 19	638 ± 247	60.0 ± 8.00	-
43	0.41 ± 0.02	0.11 ± 0.07	0.12 ± 0.04	86.1 ± 55.8	579 ± 160	120 ± 60.0	-
44	0.41 ± 0.01	0.05 ± 0.04	0.07 ± 0.05	185 ± 64.7	1638 ± 670	190 ± 116	17679 ± 2070
45	0.41 ± 0.05	0.14 ± 0.06	0.19 ± 0.13	141 ± 72.5	1503 ± 800	277 ± 240	-
77	0.16 ± 0.12	0.08 ± 0.04	0.08 ± 0.03	387 ± 140	27057 ± 5090	298 ± 250	-
78	0.19 ± 0.18	0.07 ± 0.03	0.14 ± 0.06	544 ± 231	15662 ± 1656	931 ± 688	22881 ± 1350
79	0.38 ± 0.10	0.13 ± 0.10	0.14 ± 0.07	983 ± 315	15917 ± 1626	3352 ± 1890	24606 ± 6367

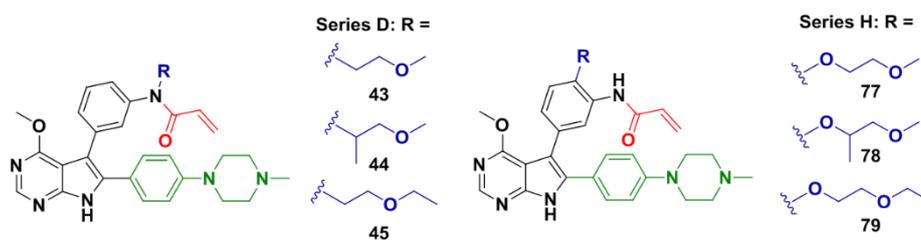
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

We postulated that the superior EC_{50} values of series D could be ascribed to a stabilising effect imparted by direct substitution of the glycol moiety on the acrylamide, resulting in a greater tolerance and stability of the chain. Consistent with previous measurements of our library of compounds, the shorter chain compounds of **43**, **44** and **77** proved to be the most efficacious, with compound **77** particularly displaying an exemplary selectivity profile.

4.6.3 Further Evaluation of Optimal Irreversible Inhibitors

While the synthesised irreversible inhibitors proved to be ineffectual against the clinically relevant C797S triple mutant, their admirable activities both biochemically and against the cell lines of the L858R and L858R/T790M mutant variants, warranted further investigation. Furthermore, as AEE788 formed the core scaffold structural basis of our library and is a dual EGFR-HER2 inhibitor, we wished to determine the potential efficacy of these compounds against HER2. We therefore opted to evaluate these compounds against four additional NSCLC cell lines, consisting of two EGFR and two clinically relevant HER2-altered NSCLC mutant cells. For EGFR, we made use of A549 cells, which are characterized as adenocarcinoma of the human alveolar basal epithelial cells and the EGFR-delE746_A750 activating mutation-carrying HCC827 cells. A549 cells contain KRAS mutations and have found use in examination of metabolic processing and mechanisms of drug delivery to lung tissue.⁸ H1781 and BaF3 HER2-insYVMA are cells found in lung cancers harbouring HER2 oncogene alterations, specifically insertion mutations within the G775_G776 region of exon 20. Ba/F3 cells are an increasingly popular tool and have gained increased use in kinase drug discovery, being adapted to high-throughput assay formats for candidate profiling. Furthermore, these cells have shown promise in predicting clinical resistance elicited by point mutations which interfere with inhibitor binding, providing invaluable data for drug development.⁹ The results of these assays may be seen in **Table 4.8** below.

Table 4.8: Evaluation of the irreversible inhibitors **43** – **45** and **77** – **79** from series D and H respectively against NSCLC and HER2 cell lines.



Compound	EGFR CTG EC_{50} [nM]			
	A549	HCC827	H1781	BaF3 HER2- insYVMA
Osimertinib	1826 ± 259	< 14	384 ± 192	364 ± 320
43	834 ± 130	< 14	341 ± 121	260 ± 72.8
44	1169 ± 131	< 14	371 ± 82.4	615 ± 531
45	1663 ± 364	< 14	548 ± 100	539 ± 288
77	955 ± 74.8	< 14	72.0 ± 2.64	340 ± 194
78	5854 ± 3656	< 14	115 ± 61.8	338 ± 146
79	1769 ± 1001	18.1 ± 7.16	83.0 ± 2.46	1102 ± 501

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Examination of the supplementary EGFR A549 cellular results once again revealed the distinctive trend of more substituted and longer chain glycol functionalities resulting in increased EC_{50} values. All compounds of series D and compound **77** exhibited superior activity compared to osimertinib, with compound **43** and **77** reaching up to 2-fold greater inhibition of the KRAS mutant cells. The resolution limit of 14 nM was reached with all but one compound when measured against the HCC287 cell line, displaying their potency against the single activating mutant. This was encouraging to us, as it showed that in nearly all cells, excluding the PC9/T790M/C797S triple mutant, the glycol chain moieties seemed to be well facilitated.

We were gratifyingly met with appealing activity profiles for the compounds of series H upon inspection of the HER2 cellular evaluation. While series D showed similar potency to osimertinib, compounds **77** – **79** all presented superior EC_{50} values close to 100 nM. For compound **77**, this equated to a 5-fold greater affinity against the H1781 HER2 mutant cells in comparison to osimertinib and lead compound **1**, with an EC_{50} value of 357 nM (not displayed). Although the pronounced gain in activity was not carried over to the BaF3 cells, good EC_{50} values were maintained for compounds **43**, **77** and **78**, highlighting their resilience towards potential resistance mutations. We correlated the compounds increased HER2 activity to the core scaffolds native dual HER2 inhibitory properties and due to TAK-285 and SYR127063, the molecules we initially drew inspiration from for glycol chain derivatisation, being dual EGFR/HER2 inhibitors as well. We postulated that the incorporation of glycol chain moieties is most beneficial in targeting HER2, thus exhibiting the potential for further development of these compounds as HER2 inhibitors.

4.7 Crystal Structure of Compound **77**

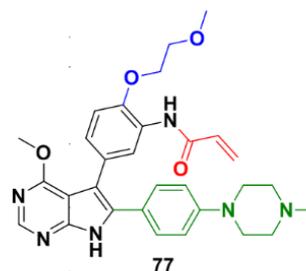
Taking into consideration the entirety of all inhibitors synthesised in our combinatorial library, compound **77** was shown to be consistently efficacious whilst maintaining an attractive selectivity profile. These results are summarized on the next page and notably include excellent activity against both the EGFR-L858R/T790M and HER2 mutant variants. For this reason, compound **77** was selected to attempt the formation of an enzyme-inhibitor complex, and to solve its subsequent crystal structure. This could provide valuable insight into the binding mode of the pyrrolopyrimidine scaffold, such as whether it undergoes covalent bond formation and illustrating any contributing interactions from the *ortho*-substituted glycol moiety.

Although the research group of Professor Rauh has succeeded in solving a series of complex crystal structures for the drug resistant mutant EGFR-T790M, compound **77** did not furnish crystals suitable for the collection of high-resolution diffraction data. Therefore, we decided to solve the structures in complex with the engineered cSrc-T338M/S345C enzyme, which has previously been successfully used as a surrogate for the EGFR-T790M mutant.^{10, 11} In this engineered c-Src enzyme, Thr338 is mutated to a methionine residue and Ser345 is mutated into a cysteine residue, corresponding to Met790 and Cys797 respectively. This composition mimics the EGFR-T790M mutant, with a bulky methionine group providing the characteristic steric interference at the gatekeeper position and the nucleophilic cysteine allowing for potential covalent modification. The inhibitor was therefore co-crystallised with cSrc-T338M/S345C, illustrated in **Figure 4.11**, through incubation with an excess of compound **77**, allowing for enzyme-inhibitor complexation prior to crystallisation, followed by crystal growth using the hanging-drop method.

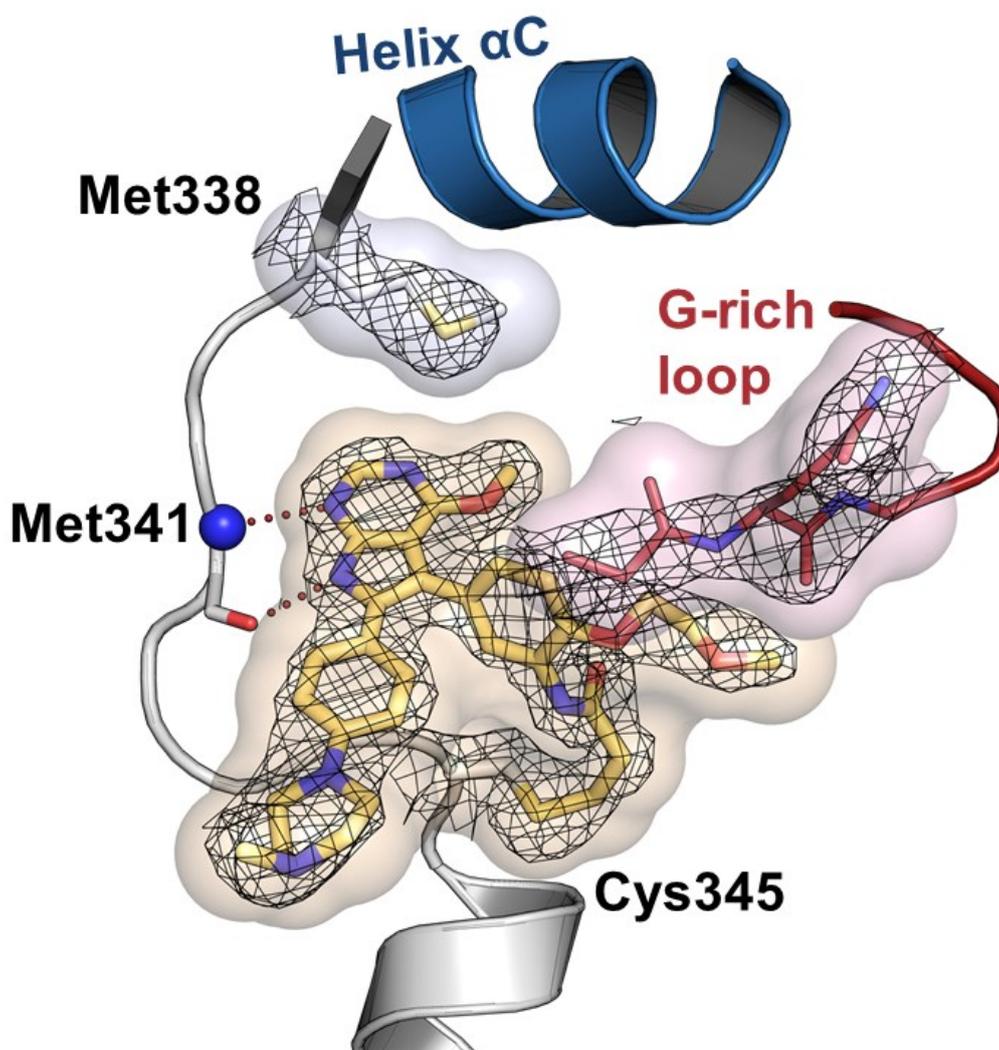
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Table 4.9: Summary of the selectivity and activity profile for compound 77.

Compound	EGFR HTRF IC ₅₀ [nM]		
	wild-type	L858R	L858R/ T790M
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40
77	0.16 ± 0.12	0.08 ± 0.04	0.08 ± 0.03



Compound	EGFR CTG EC ₅₀ [nM]					BaF3 HER2- insYVMA
	A431	A549	HCC827	H1975	H1781	
Osimertinib	756 ± 340	1826 ± 259	< 14	16.0 ± 5.00	384 ± 192	364 ± 320
77	27057 ± 5090	955 ± 74.8	< 14	298 ± 250	72.0 ± 2.64	340 ± 194

**Figure 4.11:** Co-crystal structure of compound 77 with the model cSrc-T338M/S345C system.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

The refined structure of compound **77** within the cSrc-T338M/S345C enzyme reveals a binding mode comparable to that of lead compound **1**, with the pyrrolopyrimidine heterocycle forming bidentate hydrogen bonds with Met341. It is important to note that, the hinge region Met341 residue of cSrc-T338M/S345C corresponds to Met793 of the EGFR-T790M protein kinase. Covalent bond formation between the acrylamide and Cys345 was confirmed by clear cluster of electron density in this region, as shown in **Figure 4.11**, establishing this compound as an irreversible inhibitor. The *ortho*-substituted glycol chain moiety can be seen extended towards the glycine-rich loop (**Figure 4.11**, red), potentially interacting with the amino acid residues. We believe it is these Van der Waals forces that account for the increased biochemical potency of compound **77**, over the lead compound **1**. Owing to its proximity, it is evident that an increase in the length of the glycol chain could cause steric clashes with the glycine-rich loop. Furthermore, this could lead to an unfavourable inhibitor orientation within the active site, reducing binding affinity and in a worst-case scenario, preventing covalent bond formation. This substantiates the general trend of a loss in activity associated with the longer chain glycol derivatives synthesized.

The methyl-piperazine solubilising group protrudes outwards into the solvent exposed end, anchoring the inhibitor within the binding cleft. We believe the presence of the solubilising group ensures proper reversible binding of the pyrrolopyrimidine scaffold with the hinge region, corroborating the marked loss in activity in compounds where the group is absent. Lastly, we see the pyrimidine methoxy group being able to avoid the steric clash with the gatekeeper Met338 residue, which corresponds to the Met790 residue in the EGFR-T790M mutant. Notably, the binding mode of compound **77** reveals that this position may serve as a platform for further derivatisation. Investigating the use of longer alkyl chains, which may also act as linkers for groups able to fill the conserved hydrophobic specificity pocket found in most protein kinases, might therefore be an interesting prospect.

4.8 Conclusions

The body of work described in this chapter set out to optimise and explore the effects of glycol chain derivatisation, alongside the presence or absence of the optimised methyl piperazine solubilising group and acrylamide moiety, on lead compound **1**. In doing this, we wished to gain efficacy and selectivity against the T790M and C797S mutant, by optimizing the compounds as irreversible and reversible inhibitors respectively. Inspired by the dual EGFR/HER2 inhibitors of TAK-285 and SYR127063, various chain lengths, ether functionalities and substitution positions were investigated. This led to the synthesis of a combinatorial library comprising of 26 final compounds in 8 different series (A-H). Numerous reactions were optimized throughout the synthesis, including the Suzuki-Miyaura coupling, aromatic nitro reduction and aniline-derivative monoalkylation reactions, providing products in excellent yields and the viability of a multigram synthesis.

The final compounds were evaluated against multiple mutant variants of EGFR, both biochemically and in a cellular setting. From these results, specific trends could be identified, providing beneficial information on the importance of specific functionalities in the further development of lead compound **1**. Comparison of compounds with the methyl piperazine solubilising group being present or absent showed a large disparity in activity.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

We believe that the solubilising group acts as an anchor for the rest of the molecule, ensuring proper reversible binding of the pyrrolopyrimidine core with the hinge region and proving indispensable for further optimization of these compounds.

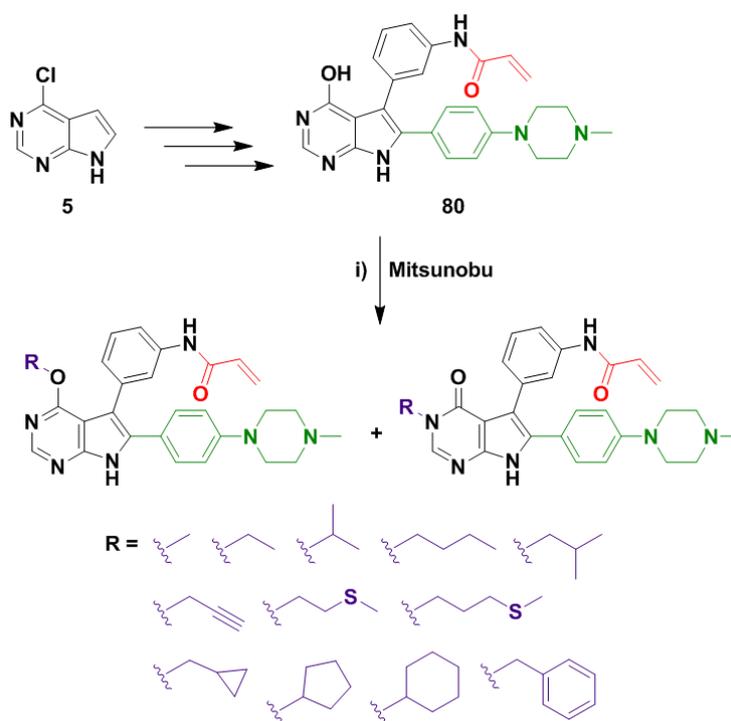
By excluding the electrophilic Michael acceptor, we were able to assess the efficacy of the synthesized compounds as reversible inhibitors, focussing particularly on their potential to inhibit the clinically relevant C797S triple mutant. As anticipated, the compounds containing the solubilising group outperformed those without by a large margin. At this stage we identified that shorter glycol chain substituents resulted in greater inhibitory efficacy, a trend which was largely preserved in future observations. Both the directly and *ortho*-substituted compounds of series C and G showed remarkable biochemical efficacy against the C797S triple mutant, with compound **40** displaying an IC₅₀ value of 1.50 nM, a 75-fold increase in activity over osimertinib. We hypothesize this to be a consequence of Van Der Waals interactions not only with the glycine-rich loop, but also with the mutated serine residue itself. However, this remarkable efficiency was diminished in the cellular evaluation against the PC9/T790M/C797S triple mutant, with compound **40** retaining a 2-fold advantage in activity over osimertinib. This was the first indication of intolerance towards the glycol chain moiety within a cell. We thus believe there is potential for the inclusion of these glycol chain moieties in generation of inhibitors specifically targeting the C797S triple mutant.

Analysis of the Michael acceptor-derivatised irreversible inhibitors revealed a recurrence of the afore-mentioned trends. Both directly and *ortho*-substituted compounds which contained the solubilising group showed exceptional biochemical activity against the L858R and L858R/T790M mutant variants, with the shorter chain glycol derivatives advancing to picomolar IC₅₀ values. We proposed that the methyl substituent of the 1-methoxypropan-2-yl glycol chain substituent could interact and promote a favourable orientation for the proximal acryl amide electrophile to undergo covalent bond formation with the cysteine residue. Disappointingly, these values were not maintained in the cellular assays, highlighting the poor pharmacokinetic properties of the glycol chains. Whilst we think further pursuit of these glycol moieties unprofitable, investigation into more stable and tolerable chain groups that can still interact with the glycine-rich loop could yield beneficial results in the development of inhibitors for the EGFR-L858R/T790M double mutant.

Owing to the dual EGFR/HER2 character of the original AEE788 core scaffold, we screened the most efficacious irreversible inhibitors against two HER2 cell lines. Encouragingly, these compounds showed good inhibitory activity, with compound **77** displaying outstanding efficacy. We believe the improved tolerance of the glycol chain and potent EC₅₀ values of compound **77** merit the further exploration and development of these compounds as dual EGFR/HER2 inhibitors.

Furthermore, we were able to obtain a crystal structure of compound **77** in the cSrc-T338M/S345C surrogate. The refined structure confirmed covalent bond formation with Cys345 and reinforced our proposals regarding the trends witnessed in the biochemical and cellular evaluation. These included visual evidence of a shorter glycol chain proving more efficacious and the importance of the solubilising group in anchoring the molecule for optimal reversible binding.

4.9 Ongoing and Future Work



Scheme 4.10: Schematic representation for the synthesis of compound **80**, which was shown to undergo a Mitsunobu reaction and afford the library of constitutional isomers. Reagents and Conditions: i) Alcohol substrate (5 equiv.), PPh₃ (5 equiv.), DIAD (5 equiv.), THF, ultrasonication at 40 °C, 6 h.

As the work undertaken in this chapter was accomplished at the Technische Universität Dortmund, research surrounding this scaffold and class of inhibitor is ongoing within the group of Prof. Rauh. We will therefore discuss the developments which have arisen post-completion of our research venture and comment on potential future endeavours. Based on the information gathered from the crystal structure of compound **77** within the cSrc-T338M/S345C enzyme, the group deduced that the most promising point for further derivatisation would be at the 4-methoxypyrimidine position. As seen in **Figure 4.11**, the pyrimidine methoxy was small enough to avoid a steric clash with the mutated methionine gatekeeper residue. This prompted the investigation of larger substituents at this position, with the potential to increase binding affinity through space filling and reversible interactions. The group was able to synthesise 4-hydroxypyrrolopyrimidine compound **80**, which was followed by a Mitsunobu reaction using ethanol and the conditions described in **Scheme 4.10**. Use of LC/MS analysis revealed the formation of two products with different retention times, but equivalent molecular mass, indicating the presence of two isomers. Subsequent separation and structural determination using 2D NMR spectroscopy unveiled the two compounds as the *O*4-substituted pyrimidine and *N*3-substituted pyrimidin-4-one (**Scheme 4.10**). This was a compelling result, sparking the reaction of numerous linear and cyclic alcohols (**Scheme 4.10**, purple) under the developed Mitsunobu conditions, to furnish a library of *O*4- and *N*3-substituted compounds. By doing so, the group ascertained the importance of both the substitution position and optimal size of alkyl or cyclic moiety to be used.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

As a future consideration, we propose that these compounds could be further derivatised with the glycol chain moieties utilised in this chapter, at either the *N*- or *ortho*-position of the acrylamide. This could be beneficial in increasing their potency as reversible inhibitors against the C797S triple mutant or by enhancing their activity as irreversible inhibitors targeting the L858R/T790M double mutant, with the absence or presence of the acrylamide functionality respectively.

Owing to the promising HER2 cellular results exhibited by the irreversible compounds shown in **Table 4.9**, the group considered the potential for development of these compounds as dual EGFR/HER2 inhibitors. Within **Figure 4.12** below, a variety of potent dual EGFR/HER2 inhibitors are displayed. These include the previously discussed SYR127063, the FDA approved inhibitor lapatinib and the preclinical candidate **81** and compound **82**, both of which have demonstrated low nanomolar IC_{50} values with a high degree of selectivity for HER2.^{12, 13} Upon closer inspection of the structural components of these compounds, we see a shared substitution pattern at the 4-pyrimidine position (**Figure 4.12**, purple). These substituents are known as back-pocket binding elements and serve to increase the binding affinity of the inhibitors by occupying the hydrophobic specificity pocket located behind the gatekeeper residue. It was therefore conceived that that these back-pocket elements could prospectively be coupled at the 4-position of our synthesized pyrrolopyrimidine core scaffold with either an ether or amine linker. This could potentially enhance their binding affinity and selectivity as dual EGFR/HER2 inhibitors and could furthermore be combined with the glycol chain moieties used in this body of research, which exhibited marked inhibitory activity against HER2 cell lines. This research is ongoing and has led to the preparation of another manuscript for publication.

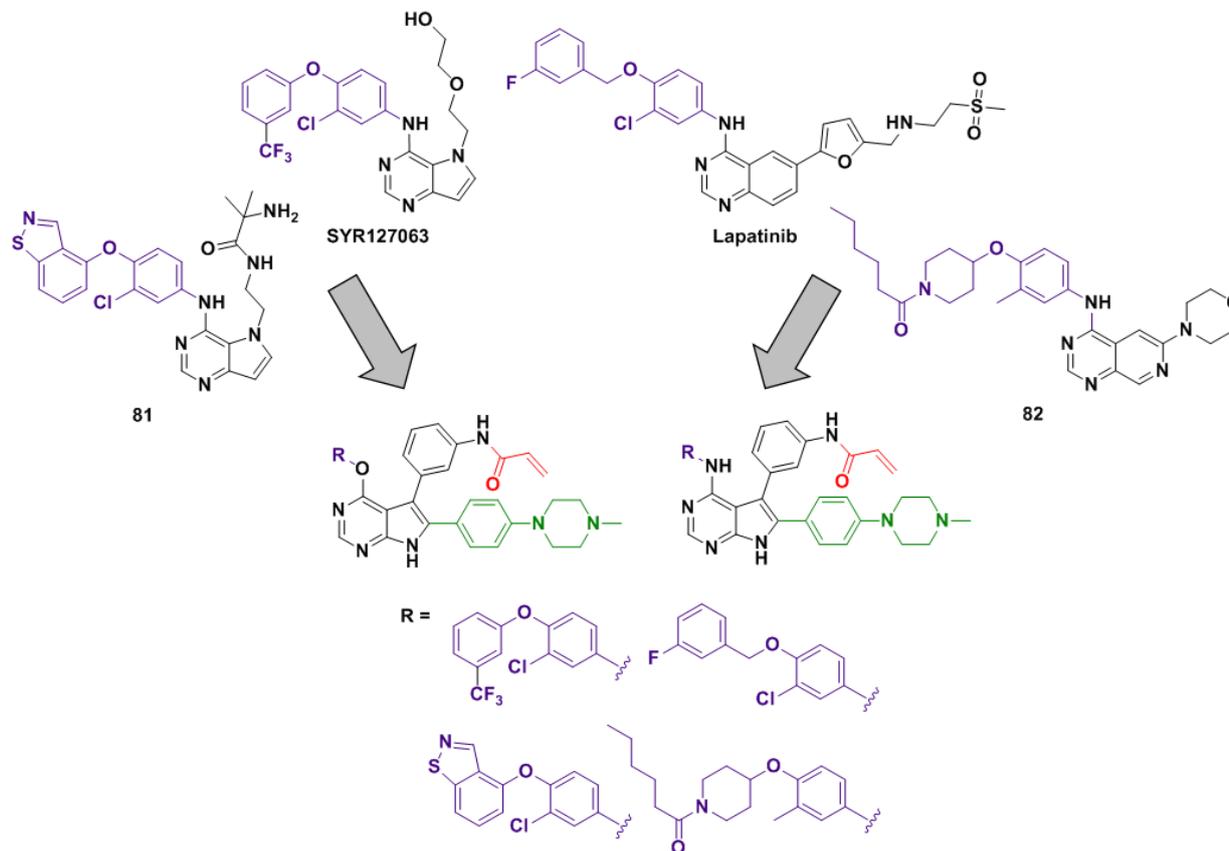
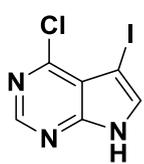


Figure 4.12: Strategy to introduce back-pocket binding elements into the synthesised pyrrolopyrimidine scaffold as potential dual EGFR/HER2 inhibitors.

4.10 Supplementary Information

4-Chloro-5-iodo-7H-pyrrolo[2,3-d]pyrimidine (6)¹⁴



A flame-dried, 2-neck round-bottomed flask was charged with 4-chloro-7H-pyrrolo[2,3-d]pyrimidine (8.00 g, 52.1 mmol, 1 equiv.), purged with Ar and dissolved in DMF (80 mL). A solution of *N*-iodosuccinimide (12.3 g, 54.7 mmol, 1.05 equiv.) in DMF (20 mL) was then added dropwise over a period of 15 min and the reaction flask was covered in foil. The reaction mixture was then allowed to stir at rt for 16 h or until complete consumption of the starting material as indicated by TLC, during which time a deep orange colour developed. After addition of a sufficient amount of 10 mol% sodium thiosulphate so that the solution became clear, the reaction mixture was diluted with H₂O (30 mL) and EtOAc (60 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 × 40 mL) and the combined organic layers were then washed with a saturated solution of brine (3 × 40 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, the residue was dried under vacuum to afford compound **6** as an off-white solid (14.1 g, 50.5 mmol, 97%) with no further purification required.

¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.¹⁴

Rf: 0.22 (50% EtOAc in PE); ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.93 (d, *J* = 2.2 Hz, 1H, ArH), 8.59 (s, 1H, ArH), 12.95 (s, 1H, ArNH) ppm; ¹³C NMR (126 MHz, DMSO-*d*₆) δ 51.7, 115.8, 133.9, 150.5, 150.8, 151.5 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₆H₄N₃ICl: 279.9133; found: 279.9132.

4-Chloro-5-iodo-7-([2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidine (7)¹⁵

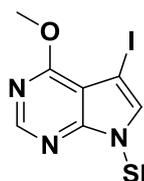


A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in oil, 1.20 g, 30.1 mmol, 1.17 equiv.), purged with Ar and to this was added DMF (90 mL). The resulting suspension was cooled to -10 °C (acetone/ice) and to this was added a solution of 4-chloro-5-iodo-7H-pyrrolo[2,3-d]pyrimidine (7.21 g, 25.8 mmol, 1 equiv.) in DMF (15 mL) dropwise over 15 min. The solution was then allowed to warm to rt and stirred for 1 h. After recooling the solution to -10 °C (acetone/ice), a solution of 2-(trimethylsilyl)ethoxymethyl chloride (5.29 g, 31.7 mmol, 1.23 equiv.) in DMF (15 mL) was added dropwise over 15 min and the solution turned yellow in colour. The reaction mixture was then allowed to warm to rt and stirred for 16 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H₂O (50 mL), resulting in the formation of a white precipitate. The precipitate was collected by filtration, washed with H₂O (3 × 30 mL) and dried under vacuum to afford compound **7** as white crystals (8.03 g, 19.6 mmol, 76%) with no further purification required.

¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.¹⁶

Rf: 0.45 (10% EtOAc in PE); ¹H NMR (500 MHz, DMSO-*d*₆) δ -0.11 (s, 9H, CH₂Si(CH₃)₃), 0.81 (t, *J* = 8.0 Hz, 2H, CH₂CH₂Si), 3.50 (t, *J* = 8.0 Hz, 2H, OCH₂CH₂), 5.60 (s, 2H, NCH₂O), 8.12 (s, 1H, ArH), 8.68 (s, 1H, ArH) ppm; ¹³C NMR (126 MHz, DMSO-*d*₆) δ -1.4, 17.1, 52.9, 66.1, 72.9, 116.2, 136.6, 150.9, 151.1, 151.2 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₂H₁₈N₃OSiCl: 409.9952; found: 409.9950.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

5-Iodo-4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (8)¹⁷

A 2-neck round-bottomed flask was charged with 4-chloro-5-iodo-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (3.35 g, 8.18 mmol, 1 equiv.), potassium carbonate (2.49 g, 18.0 mmol, 2.2 equiv.), purged with Ar and dissolved in MeOH (40 mL). The solution was then allowed to stir at rt for 16 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, the residue was diluted with H₂O (60 mL) and EtOAc (60 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 × 40 mL) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography with elution gradient of 0 – 20% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **8** as a white solid (3.05 g, 7.53 mmol, 92%).

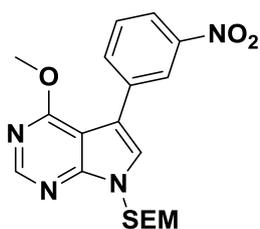
¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.¹⁶

Rf: 0.65 (15% EtOAc in PE); ¹H NMR (500 MHz, DMSO-*d*₆) δ –0.11 (s, 9H, CH₂Si(CH₃)₃), 0.80 (t, *J* = 8.0 Hz, 2H, CH₂CH₂Si), 3.48 (t, *J* = 8.0 Hz, 2H, OCH₂CH₂), 4.05 (s, 3H, ArOCH₃), 5.53 (s, 2H, NCH₂O), 7.75 (s, 1H, ArH), 8.45 (s, 1H, ArH) ppm; ¹³C NMR (126 MHz, DMSO-*d*₆) δ –1.4, 17.1, 50.9, 53.74, 65.7, 72.5, 106.3, 132.3, 151.3, 151.9, 162.3 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₃H₂₁N₃O₂Si: 406.0448; found: 406.0451.

General procedure for Suzuki-Miyaura cross coupling reaction:¹⁸

A flame dried Schlenk tube was purged with Ar and to this was added a 3:1 solution of dioxane/H₂O (5 mL/mmol aryl halide) via syringe under positive pressure Ar. The solvent was then degassed using 3 – 5 cycles of freeze-pump-thaw technique, followed by bubbling Ar under positive pressure for 10 min. A separate Schlenk tube was charged with aryl halide (1 equiv.), boronic acid/ester (1.2 equiv.), Pd(PPh₃)₄ (5 mol%), sodium carbonate (2.5 equiv.) and purged with Ar. To this was added the previously degassed solvent via syringe, under positive pressure Ar. The reaction flask was sealed, heated to 80 °C and allowed to stir for 16 h or until complete consumption of the starting material as indicated by TLC, during which time the reaction mixture developed a red-brown colour. After allowing to cool to rt, the reaction mixture was filtered through a plug of celite, diluted with H₂O (10 mL/mmol aryl halide) and EtOAc (20 mL/mmol aryl halide) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 × 20 mL/mmol aryl halide) and the combined organic layers were then washed with a saturated solution of brine (10 mL/mmol aryl halide), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

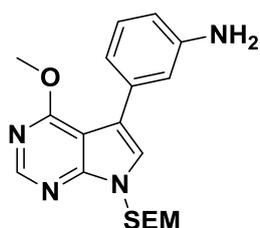
4-Methoxy-5-(3-nitrophenyl)-7-([2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidine (9)

Compound **9** was prepared according to the general procedure for Suzuki-Miyaura cross coupling reaction. 5-Iodo-4-methoxy-7-([2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidine (405 mg, 1.00 mmol) and (3-nitrophenyl)boronic acid (200 mg, 1.20 mmol) were coupled and purification was achieved using flash column chromatography with elution gradient of 0 – 20% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **9** as a yellow semi-solid (368 mg, 0.918 mmol, 92%).

Rf: 0.37 (15% EtOAc in PE); ^1H NMR (500 MHz, DMSO- d_6) δ -0.10 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.84 (t, $J = 7.9$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.56 (t, $J = 8.0$ Hz, 2H, OCH_2CH_2), 4.05 (s, 3H, ArOCH_3), 5.63 (s, 2H, NCH_2O), 7.68 (dd [app. t], $J = 8.0$ Hz, 1H, ArH), 8.07 (s, 1H, ArH), 8.13 – 8.10 (m, 2H, $2 \times \text{ArH}$), 8.59 (dd [app. t], $J = 2.0$ Hz, 1H, ArH), 8.51 (s, 1H, ArH) ppm; ^{13}C NMR (500 MHz, DMSO- d_6) δ -1.4, 17.1, 53.8, 65.9, 72.8, 102.2, 113.8, 121.0, 122.7, 126.9, 129.7, 134.3, 135.3, 147.9, 151.3, 152.8, 162.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{25}\text{N}_4\text{O}_4\text{Si}$: 401.1645; found: 401.1642.

General procedure for aromatic nitro derivative reduction:¹⁹

A 2-neck round-bottomed flask was charged with aromatic nitro derivative (1 equiv.), iron (5 equiv.) and ammonium chloride (1 equiv.) and to this was added a 1:1 solution of EtOH/ H_2O (5 mL/mmol aromatic nitro derivative). The resulting suspension were place under ultrasonic irradiation at 60 °C for 6 h or until complete consumption of the starting material as indicated by TLC, during which time the reaction mixture developed a dark-brown colour. After allowing to cool to rt, the reaction mixture was filtered through a plug of celite, washed with EtOH (5 mL/mmol aromatic nitro derivative) and the solvent was removed *in vacuo*. The residue was then diluted with DCM (10 mL/mmol aromatic nitro derivative) and H_2O (10 mL/mmol aromatic nitro derivative) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×10 mL/mmol aromatic nitro derivative) and the combined organic layers were then washed with a saturated solution of brine (10 mL/mmol aryl halide), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

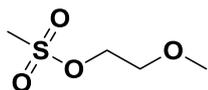
3-(4-Methoxy-7-([2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (10)

Compound **10** was prepared was according to the general procedure for aromatic nitro derivative reduction. 4-Methoxy-5-(3-nitrophenyl)-7-([2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidine (1.60 g, 4.00 mmol) was reduced and purification was achieved using flash column chromatography with elution gradient of 20 – 80% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **10** as a light-brown

semi-solid (1.29 g, 3.48 mmol, 87%).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Rf: 0.44 (50% EtOAc in PE); ^1H NMR (400 MHz, CDCl_3) δ -0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.92 (t, $J = 7.9$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.56 (t, $J = 7.8$ Hz, 2H, OCH_2CH_2), 3.87 (br s, 2H, ArNH_2), 4.07 (s, 3H, ArOCH_3), 5.62 (s, 2H, NCH_2O), 6.65 (ddd, $J = 7.9, 2.2, 0.8$ Hz, 1H, ArH), 7.00 (t, $J = 1.9$ Hz, 1H, ArH), 7.06 (dt, $J = 7.8, 1.1$ Hz, 1H, ArH), 7.19 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.24 (s, 1H, ArH), 8.49 (s, 1H, ArH) ppm; ^{13}C NMR (101 MHz, CDCl_3) δ -1.4, 17.8, 53.7, 66.5, 73.1, 103.4, 113.8, 115.8, 118.0, 119.5, 123.7, 129.1, 134.9, 146.0, 151.3, 153.0, 163.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{27}\text{N}_4\text{O}_2\text{Si}$: 371.1903; found: 371.1902.

2-Methoxyethyl methanesulfonate (12)²⁰

A flame-dried, 2-neck round-bottom flask was charged with 2-methoxyethanol (2.00 g, 26.3 mmol, 1 equiv.), purged with Ar and dissolved in DCM (50 mL). To this was added Et_3N (4.40 mL, 31.5 mmol, 1.2 equiv.) and the solution was cooled to -10 °C (acetone/ice). This was followed by the dropwise addition of a solution of methanesulfonyl chloride (3.31 g, 28.9 mmol, 1.1 equiv.) in DCM (25 mL) and the reaction mixture was then allowed to warm to rt and stirred for 12 h. The reaction mixture was then quenched and diluted with H_2O (30 mL) and DCM (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×40 mL) and the combined organic layers were then washed with H_2O (2×50 mL), a solution of 0.1 M HCl (50 mL), a saturated solution of brine (50 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved by kugelrohr distillation. Fractions with boiling point between 135 °C and 145 °C at ~ 20 mbar were collected to afford compound **12** as a clear oil (3.57 g, 23.1 mmol, 88%).

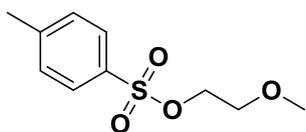
^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.²¹

^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 3.17 (s, 3H, OSO_2CH_3), 3.28 (s, 3H, CH_2OCH_3), 3.61 – 3.56 (m, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.33 – 4.28 (m, 2H, $\text{OCH}_2\text{CH}_2\text{O}$) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 36.7, 58.1, 69.4, 69.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{27}\text{N}_4\text{O}_2\text{Si}$: 371.1903; found: 371.1908.

General procedure for tosylation of glycol ethers:²²

A flame-dried, 2-neck round-bottom flask was charged with glycol ether (1 equiv.), purged with Ar and dissolved in DCM (2 mL/mmol glycol ether). To this was added Et_3N (1.2 equiv.) and the solution was cooled to -10 °C (acetone/ice). This was followed by the dropwise addition of a solution of *p*-toluenesulfonyl chloride (1.1 equiv.) in DCM (1 mL/mmol glycol ether) and the reaction mixture was then allowed to warm to rt and stirred for 12 h. The reaction mixture was then quenched and diluted with H_2O (1 mL/mmol glycol ether) and DCM (1 mL/mmol glycol ether) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×1 mL/mmol glycol ether) and the combined organic layers were then washed with H_2O (2×2 mL/mmol glycol ether), a solution of 0.1 M HCl (2 mL/mmol glycol ether), a saturated solution of brine (2 mL/mmol glycol ether), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

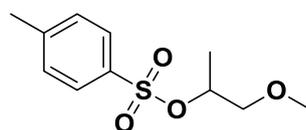
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

2-Methoxyethyl 4-methylbenzenesulfonate (13)

Compound **13** was prepared according to the general procedure for tosylation of glycol ethers. 2-Methoxyethanol (1.00 g, 13.1 mmol) was tosylated and purification was achieved using flash column chromatography with elution gradient of 10 – 40% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **13** as a clear oil (2.78 g, 12.1 mmol, 92%).

¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.²³

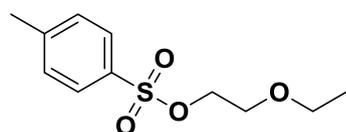
Rf: 0.14 (10% EtOAc in PE); ¹H NMR (500 MHz, DMSO-*d*₆) δ 2.41 (s, 3H, ArCH₃), 3.17 (s, 3H, CH₂OCH₃), 3.48 (t, *J* = 4.4 Hz, 2H, OCH₂CH₂O), 4.11 (t, *J* = 4.4 Hz, 2H, OCH₂CH₂O), 7.47 (d, *J* = 8.2 Hz, 2H, 2 × ArH), 7.79 (d, *J* = 8.2 Hz, 2H, 2 × ArH) ppm; ¹³C NMR (126 MHz, DMSO-*d*₆) δ 21.1, 57.9, 69.3, 69.8, 127.7, 130.2, 132.4, 145.0 ppm; HRMS-TOF MS ESI+: *m/z* [M+Na]⁺ calculated for C₁₀H₁₄NaO₄S: 253.0505; found: 253.0499.

1-Methoxypropan-2-yl 4-methylbenzenesulfonate (14)

Compound **14** was prepared according to the general procedure for tosylation of glycol ethers. 1-Methoxypropan-2-ol (1.00 g, 11.1 mmol) was tosylated and purification was achieved using flash column chromatography with elution gradient of 10 – 40% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **14** as a clear oil (2.47 g, 10.1 mmol, 91%).

¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.²⁴

Rf: 0.18 (10% EtOAc in PE); ¹H NMR (500 MHz, CDCl₃) δ 1.23 (d, *J* = 6.8 Hz, 3H, CH₃CHCH₂), 2.41 (s, 3H, ArCH₃), 3.20 (s, 3H, CHCH₂OCH₃), 3.41 – 3.30 (m, 2H, CHCH₂OCH₃), 4.71 – 4.63 (m, 1H, TsOCH(CH₃)CH₂), 7.30 (d, *J* = 8.0 Hz, 2H, 2 × ArH), 7.77 (d, *J* = 8.0 Hz, 2H, 2 × ArH) ppm; ¹³C NMR (126 MHz, CDCl₃) δ 17.6, 21.7, 59.1, 74.7, 78.0, 127.9, 129.7, 134.2, 144.6 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₁H₁₇O₄S: 245.0848; found: 245.0848.

2-Ethoxyethyl 4-methylbenzenesulfonate (15)

Compound **15** was prepared according to the general procedure for tosylation of glycol ethers. 3-Methoxy-1-propanol (1.00 g, 11.1 mmol) was tosylated and purification was achieved using flash column chromatography with elution gradient of 10 – 40% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **15** as a clear oil (2.44 g, 9.99 mmol, 90%).

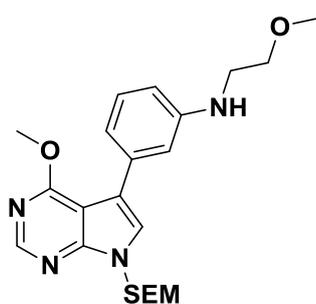
¹H NMR data collected for this compound compared well with the reported literature values.²⁵

Rf: 0.18 (10% EtOAc in PE); ¹H NMR (500 MHz, CDCl₃) δ 1.09 (t, *J* = 7.0 Hz, 3H, OCH₂CH₃), 2.40 (s, 3H, ArCH₃), 3.40 (q, *J* = 7.0 Hz, 2H, OCH₂CH₃), 3.56 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 4.11 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 7.31 (d, *J* = 8.2 Hz, 2H, 2 × ArH), 7.76 (d, *J* = 8.2 Hz, 2H, 2 × ArH) ppm; ¹³C NMR (126 MHz, CDCl₃) δ 15.0, 21.6, 66.7, 67.8, 69.4, 128.0, 129.8, 133.0, 144.9 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₁H₁₇O₄S: 245.0848; found: 245.0850.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

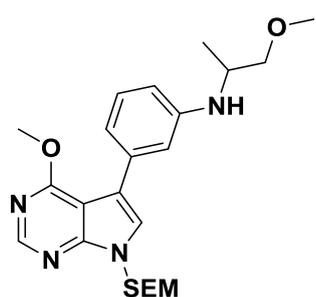
General procedure for aniline derivative monoalkylation:²⁶

A microwave reaction vial was charged with aniline derivative (1.5 equiv.), tosylated glycol ether (1 equiv.) and potassium carbonate (1.5 equiv.) and to this was added DMF (4 mL/mmol aniline derivative). The vial was purged with Ar, sealed and the reaction mixture was then heated to 80 °C and stirred for 72 h. After cooling to rt, the reaction mixture was diluted with H₂O (20 mL/mmol glycol ether) and EtOAc (20 mL/mmol glycol ether) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 × 10 mL/mmol glycol ether) and the combined organic layers were then washed with a saturated solution of brine (4 × 15 mL/mmol glycol ether), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound and excess starting aniline derivative was recovered and reused in repeat experiments.

3-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-N-(2-methoxyethyl)aniline (16)

Compound **16** was prepared according to the general procedure for aniline derivative monoalkylation. 3-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (404 mg, 1.09 mmol) was alkylated with 2-methoxyethyl 4-methylbenzenesulfonate (168 mg, 0.730 mmol) and purification was achieved using flash column chromatography with elution gradient of 10 – 80% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **16** as a yellow semi-solid (203 mg, 0.474 mmol, 65% isolated, 90% brsm).

R_f: 0.58 (50% EtOAc in PE); ¹H NMR (400 MHz, CDCl₃) δ -0.04 (s, 9H, CH₂Si(CH₃)₃), 0.94 (t, 2H, J = 8.2 Hz, OCH₂CH₂Si), 3.38 – 3.36 (m, 2H, NHCH₂CH₂O), 3.40 (s, 3H, CH₂OCH₃), 3.58 (t, J = 8.2 Hz, 2H, OCH₂CH₂Si), 3.65 (t, J = 5.2 Hz, 2H, NHCH₂CH₂O), 4.10 (s, 3H, ArOCH₃), 5.65 (s, 2H, NCH₂O), 6.64 (dd, J = 8.0, 1.5 Hz, 1H, ArH), 7.01 (dd, J = 3.8, 1.7 Hz, 1H, ArH), 7.03 (dd, J = 7.8, 1.0 Hz, 1H, ArH), 7.22 (dd [app. t], J = 7.8 Hz, 1H, ArH), 7.28 (dd [app. t], J = 4.2 Hz, 1H, ArH), 8.51 (s, 1H, ArH) ppm; ¹³C NMR (101 MHz, CDCl₃) δ -1.4, 17.8, 43.8, 53.7, 58.8, 66.5, 71.1, 73.1, 103.4, 112.2, 113.7, 118.4, 118.5, 123.7, 129.1, 134.9, 148.0, 151.3, 153.1, 163.5 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₂H₃₃N₄O₃Si: 429.2322; found: 429.2321.

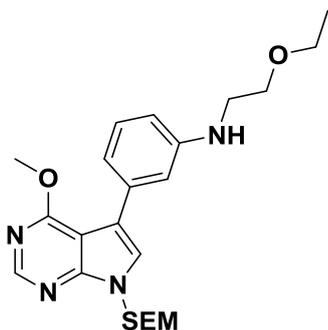
3-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-N-(1-methoxypropan-2-yl)aniline (18)

Compound **18** was prepared according to the general procedure for aniline derivative monoalkylation. 3-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (365 mg, 0.985 mmol) was alkylated with 1-methoxypropan-2-yl 4-methylbenzenesulfonate (161 mg, 0.659 mmol) and purification was achieved using flash column chromatography with elution gradient of 10 – 80% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **18** as a yellow semi-solid (161 mg, 0.364 mmol, 55% isolated, 78% brsm).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Rf: 0.74 (50% EtOAc in PE); ^1H NMR (400 MHz, CDCl_3) δ -0.05 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.93 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.28 (d, $J = 6.8$ Hz, 3H, CH_3CHCH_2), 3.38 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.52 – 3.41 (m, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.57 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.77 – 3.68 (m, 1H, CH_3CHCH_2), 4.09 (s, 3H, ArOCH_3), 5.64 (s, 2H, NCH_2O), 6.63 (dd, $J = 8.0, 1.4$ Hz, 1H, ArH), 7.03 – 6.99 (m, $2 \times \text{ArH}$), 7.21 (dd [app. t], $J = 8.0$ Hz, 1H, ArH), 7.27 (s, 1H, ArH), 8.50 (s, 1H, ArH) ppm; ^{13}C NMR (101 MHz, CDCl_3) δ -1.3, 17.8, 18.0, 48.9, 53.8, 59.2, 66.6, 73.1, 76.1, 103.5, 112.9, 114.2, 118.4, 118.5, 123.7, 129.2, 135.0, 146.9, 151.3, 153.1, 163.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{35}\text{N}_4\text{O}_3\text{Si}$: 443.2478; found: 443.2482.

N-(2-Ethoxyethyl)-3-(4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (19)



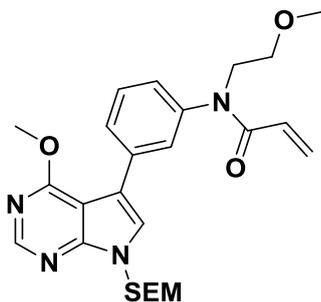
Compound **19** was prepared according to the general procedure for aniline derivative monoalkylation. 3-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (390 mg, 1.05 mmol) was alkylated with 2-ethoxyethyl 4-methylbenzenesulfonate (171 mg, 0.700 mmol) and purification was achieved using flash column chromatography with elution gradient of 10 – 80% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **19** as a yellow semi-solid (192 mg, 0.434 mmol, 62% isolated, 82% brsm).

Rf: 0.72 (50% EtOAc in PE); ^1H NMR (400 MHz, CDCl_3) δ -0.05 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.92 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.22 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 3.35 (t, $J = 5.3$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 3.59 – 3.51 (m, 4H, $\text{OCH}_2\text{CH}_2\text{Si}$ and OCH_2CH_3), 3.68 (t, $J = 5.3$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 4.09 (s, 3H, ArOCH_3), 5.64 (s, 2H, NCH_2O), 6.63 (dd, $J = 8.3, 2.1$ Hz, 1H, ArH), 7.04 – 6.99 (m, 2H, $2 \times \text{ArH}$), 7.21 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.27 (s, 1H, ArH), 8.50 (s, 1H, ArH) ppm; ^{13}C NMR (101 MHz, CDCl_3) δ -1.4, 15.3, 17.8, 44.0, 53.7, 66.5, 66.5, 68.9, 73.1, 103.5, 112.3, 113.8, 118.4, 118.5, 123.7, 129.1, 134.8, 148.0, 151.3, 153.1, 163.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{35}\text{N}_4\text{O}_3\text{Si}$: 443.2478; found: 443.2479.

General procedure for aniline derivative acylation:

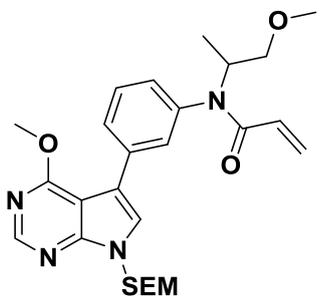
A flame-dried, 2-neck round-bottom flask was charged with aniline derivative (1 equiv.), potassium carbonate (1.5 equiv.), purged with Ar and to this was added THF (5 mL/mmol aniline derivative). The resulting suspension was cooled to -10 °C (acetone/ice) and to this was added a solution of acryloyl chloride (1.2 equiv.) in THF (5 mL/mmol aniline derivative) dropwise over 15 min. The reaction mixture was then allowed to warm to rt and stirred for 6 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (20 mL/mmol aniline derivative) and EtOAc (20 mL/mmol aniline derivative) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×10 mL/mmol aniline derivative) and the combined organic layers were then washed with a saturated solution of brine (20 mL/mmol aniline derivative), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-[3-(4-Methoxy-7-[2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidin-5-yl]phenyl]-N-(2-methoxyethyl)acrylamide (20)

Compound **20** was prepared according to the general procedure for aniline derivative acylation. 3-(4-Methoxy-7-[2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-N-(2-methoxyethyl)aniline (165 mg, 0.385 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **20** as a clear semi-solid (163 mg, 0.338 mmol, 89%).

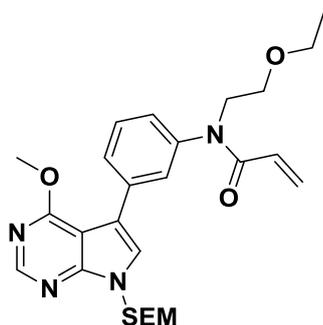
Rf: 0.29 (50% EtOAc in PE); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ -0.09 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.90 (t, $J = 8.1$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.29 (s, 3H, CH_2OCH_3), 3.54 (t, $J = 8.1$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.58 (t, $J = 6.0$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 3.98 (t, $J = 6.0$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 4.04 (s, 3H, ArOCH_3), 5.51 (d, $J = 10.5$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 5.63 (s, 2H, NCH_2O), 6.14 (dd, $J = 16.7, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.37 (d, $J = 16.7$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.12 (d, $J = 7.8$ Hz, 1H, ArH), 7.30 (s, 1H, ArH), 7.40 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.54 (s, 1H, ArH), 7.58 (d, $J = 7.8$ Hz, 1H, ArH), 8.48 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ -1.4, 17.7, 48.9, 53.9, 58.6, 66.6, 69.4, 73.1, 103.2, 116.8, 124.0, 126.6, 127.7, 127.8, 128.8, 128.9, 129.3, 135.5, 142.1, 151.4, 153.0, 163.4, 165.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{25}\text{H}_{35}\text{N}_4\text{O}_3\text{Si}$: 467.2478; found: 467.2473.

N-[3-(4-Methoxy-7-[2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidin-5-yl]phenyl]-N-(1-methoxypropan-2-yl)acrylamide (21)

Compound **21** was prepared according to the general procedure for aniline derivative acylation. 3-(4-Methoxy-7-[2-(trimethylsilyl)ethoxy]methyl)-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-N-(1-methoxypropan-2-yl)aniline (89.0 mg, 0.201 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **21** as a clear semi-solid (89.0 mg, 0.179 mmol, 89%).

Rf: 0.36 (50% EtOAc in PE); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ -0.08 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.90 (t, $J = 8.1$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.11 (d, $J = 7.0$ Hz, 3H, CH_3CHCH_2), 3.31 – 3.26 (m, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.32 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.55 (t, $J = 8.1$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 4.03 (s, 3H, ArOCH_3), 4.11 – 4.05 (m, 1H, CH_3CHCH_2), 5.46 (d, $J = 10.4$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 5.64 (s, 2H, NCH_2O), 5.96 (dd, $J = 16.7, 10.4$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.35 (dd, $J = 16.7, 1.1$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.18 – 7.01 (m, 1H, ArH), 7.30 (s, 1H, ArH), 7.42 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.55 – 7.44 (m, 1H, ArH), 7.63 (d, $J = 7.8$ Hz, 1H, ArH), 8.49 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ -1.4, 15.7, 17.8, 54.0, 58.5, 60.5, 66.7, 73.2, 73.6, 103.2, 116.9, 124.1, 127.5, 128.3, 128.7, 129.1, 129.4, 131.3, 135.2, 151.4, 153.0, 163.5, 165.8, 171.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{26}\text{H}_{37}\text{N}_4\text{O}_4\text{Si}$: 497.2584; found: 497.2588.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-(2-Ethoxyethyl)-N-[3-(4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]acrylamide (22)

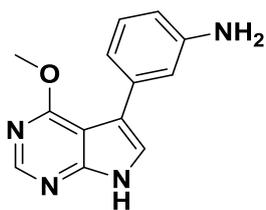
Compound **22** was prepared according to the general procedure for aniline derivative acylation. *N*-(2-Ethoxyethyl)-3-(4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (172 mg, 0.389 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **22** as a clear semi-solid (174 mg, 0.350 mmol, 90%).

Rf: 0.35 (50% EtOAc in PE); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ -0.09 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.90 (t, $J = 8.0$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.08 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 3.43 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.54 (t, $J = 8.0$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.64 (t, $J = 6.1$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 3.97 (t, $J = 6.1$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 4.04 (s, 3H, ArOCH_3), 5.51 (d, $J = 10.5$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 5.63 (s, 2H, NCH_2O), 6.15 (dd, $J = 16.7, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.37 (dd, $J = 16.8, 1.4$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.13 (d, $J = 7.8$ Hz, 1H, ArH), 7.29 (s, 1H, ArH), 7.39 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.56 (s, 1H, ArH), 7.58 (d, $J = 8.0$ Hz, 1H, ArH), 8.48 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ -1.4, 15.2, 17.7, 53.9, 60.4, 66.2, 66.6, 67.4, 73.1, 103.2, 116.9, 124.0, 126.6, 127.6, 127.7, 128.9, 128.9, 129.2, 135.4, 142.3, 151.4, 153.0, 163.4, 165.7, 171.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{26}\text{H}_{37}\text{N}_4\text{O}_4\text{Si}$: 497.2584; found: 497.2587.

General procedure for N-SEM deprotection:²⁷

A round-bottom flask was charged with *N*-SEM protected pyrrolopyrimidine derivative (1 equiv.) and dissolved in DCM (15 mL/mmol pyrrolopyrimidine derivative). To this was added trifluoroacetic acid (5 mL/mmol pyrrolopyrimidine derivative) dropwise and the reaction flask was stoppered and allowed to stir for 6 h at rt, during which time the reaction mixture developed a brown colour. After removal of the solvent *in vacuo*, the residue was dissolved in THF (10 mL/mmol pyrrolopyrimidine derivative), followed by the slow addition of a solution of 1 M KOH (10 mL/mmol pyrrolopyrimidine derivative). The reaction mixture was then allowed to stir for a further 4 h, during which time the solution became clear. The reaction mixture was then neutralised by the slow addition of a solution of 1 M HCl (10 mL/mmol pyrrolopyrimidine derivative), diluted with H_2O (20 mL/mmol pyrrolopyrimidine derivative) and DCM (30 mL mmol pyrrolopyrimidine derivative) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL/mmol pyrrolopyrimidine derivative) and the combined organic layers were then washed with a saturated solution of brine (20 mL/mmol pyrrolopyrimidine derivative), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

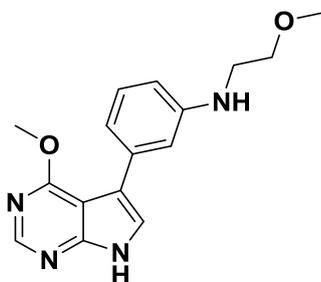
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

3-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (23)

Compound **23** was prepared according to the general procedure for *N*-SEM deprotection.

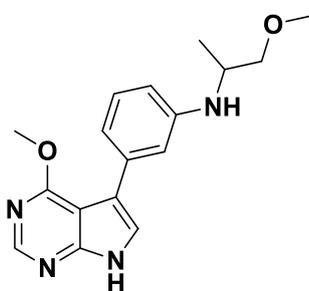
3-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (126 mg, 0.340 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **23** as a white solid (64.0 mg, 0.266 mmol, 78%).

Rf: 0.41 (5% MeOH in DCM); ^1H NMR (500 MHz, DMSO- d_6) δ 4.03 (s, 3H, ArOCH₃), 6.81 (d, J = 7.6 Hz, 1H, ArH), 7.26 – 7.17 (m, 3H, 3 \times ArH), 7.54 (s, 1H, ArH), 8.41 (s, 1H, ArH), 12.29 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, DMSO- d_6) δ 53.5, 102.0, 115.7, 118.0, 118.4, 121.2, 122.4, 124.8, 128.9, 135.3, 150.6 ppm; HRMS-TOF MS ESI+: m/z [M+H]⁺ calculated for C₁₃H₁₃N₄O: 241.1089; found: 241.1092.

3-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-N-(2-methoxyethyl)aniline (24)

Compound **24** was prepared according to the general procedure for *N*-SEM deprotection. 3-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-*N*-(2-methoxyethyl)aniline (132 mg, 0.308 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **24** as a white solid (75.0 mg, 0.251 mmol, 81%).

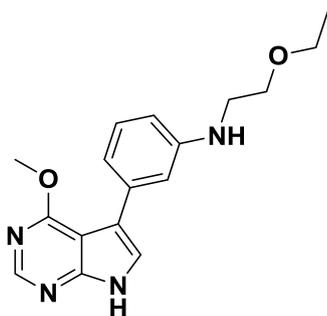
Rf: 0.36 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl₃) δ 3.44 (t, J = 5.1 Hz, 2H, NCH₂CH₂O), 3.48 (s, 3H, CH₂OCH₃), 3.73 (t, J = 5.1 Hz, 2H, NCH₂CH₂O), 4.15 (s, 3H, ArOCH₃), 4.52 (br s, 1H, ArNHCH₂), 6.71 (d, J = 6.8 Hz, 1H, ArH), 7.07 (s, 1H, ArH), 7.10 (d, J = 7.6 Hz, 1H, ArH), 7.29 (d, J = 7.8 Hz, 1H, ArH), 7.32 (s, 1H, ArH), 8.57 (s, 1H, ArH), 12.45 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl₃) δ 43.7, 53.7, 58.7, 71.0, 103.2, 111.9, 113.7, 117.7, 118.4, 121.2, 129.0, 135.2, 147.9, 150.3, 152.9, 163.5 ppm; HRMS-TOF MS ESI+: m/z [M+H]⁺ calculated for C₁₆H₁₉N₄O₂: 299.1508; found: 299.1505.

3-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-N-(1-methoxypropan-2-yl)aniline (25)

Compound **25** was prepared according to the general procedure for *N*-SEM deprotection. 3-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-*N*-(1-methoxypropan-2-yl)aniline (111 mg, 0.251 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **25** as a white solid (64.0 mg, 0.205 mmol, 82%).

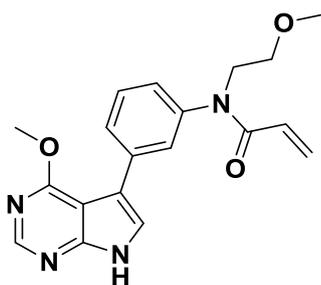
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Rf: 0.35 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 1.31 (d, $J = 6.6$ Hz, 3H, CH_3CHCH_2), 3.39 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.54 – 3.43 (m, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.79 – 3.71 (m, 1H, CH_3CHCH_2), 4.10 (s, 3H, ArOCH_3), 4.35 (br s, 1H, ArNHCH), 6.68 (d, $J = 7.8$ Hz, 1H, ArH), 7.09 – 7.03 (m, 2H, $2 \times \text{ArH}$), 7.23 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.30 (s, 1H, ArH), 8.52 (s, 1H, ArH), 11.62 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 17.8, 49.4, 53.9, 59.2, 75.8, 103.4, 113.3, 114.9, 117.9, 119.3, 121.3, 129.2, 135.4, 146.1, 150.6, 153.1, 163.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{17}\text{H}_{21}\text{N}_4\text{O}_2$: 313.1664; found: 313.1665.

N-(2-Ethoxyethyl)-3-(4-methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (26)

Compound **26** was prepared according to the general procedure for *N*-SEM deprotection. *N*-(2-Ethoxyethyl)-3-(4-methoxy-7- $\{[2-(\text{trimethylsilyl})\text{ethoxy}]\text{methyl}\}$ -7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (112 mg, 0.253 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **26** as a white solid (66.0 mg, 0.211 mmol, 83%).

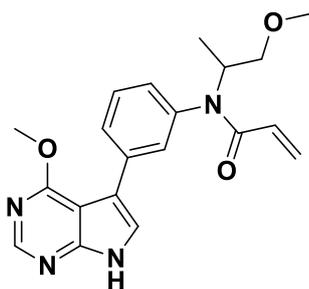
Rf: 0.36 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 1.23 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 3.37 (t, $J = 5.2$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 3.55 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.70 (t, $J = 5.2$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{O}$), 4.09 (s, 3H, ArOCH_3), 4.44 (br s, 1H, ArNHCH_2), 6.65 (d, $J = 7.3$ Hz, 1H, ArH), 7.03 – 6.96 (m, 1H, ArH), 7.05 (d, $J = 8.0$ Hz, 1H, ArH), 7.23 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.29 (s, 1H, ArH), 8.51 (s, 1H, ArH), 12.04 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 15.3, 44.1, 53.8, 66.5, 68.9, 103.4, 112.2, 114.0, 117.9, 118.7, 121.2, 129.1, 135.2, 147.8, 150.5, 153.0, 163.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{17}\text{H}_{21}\text{N}_4\text{O}_2$: 313.1665; found: 313.1666.

N-[3-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]-N-(2-methoxyethyl)acrylamide (27)

Compound **27** was prepared according to the general procedure for *N*-SEM deprotection. *N*-[3-(4-Methoxy-7- $\{[2-(\text{trimethylsilyl})\text{ethoxy}]\text{methyl}\}$ -7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]-*N*-(2-methoxyethyl)acrylamide (120 mg, 0.249 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **27** as a white solid (75.0 mg, 0.213 mmol, 86%).

Rf: 0.41 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 3.31 (s, 3H, CH_3OCH_3), 3.61 (t, $J = 6.0$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 4.02 (t, $J = 6.0$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 4.06 (s, 3H, ArOCH_3), 5.54 (d, $J = 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.18 (dd, $J = 16.6, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.41 (d, $J = 16.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.14 (d, $J = 7.8$ Hz, 1H, ArH), 7.37 (s, 1H, ArH), 7.42 (dd [app. t], $J = 7.9$ Hz, 1H, ArH), 7.58 (s, 1H, ArH), 7.62 (d, $J = 7.7$ Hz, 1H, ArH), 8.51 (s, 1H, ArH), 12.40 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 48.9, 53.9, 58.7, 69.4, 103.2, 116.3, 121.8, 124.4, 126.4, 127.8, 127.9, 128.8, 129.3, 135.9, 142.0, 150.7, 153.1, 163.6, 165.9 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{21}\text{N}_4\text{O}_3$: 353.1614; found: 353.1612.

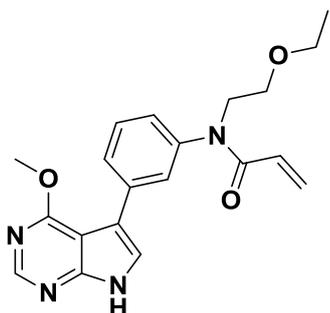
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-[3-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]-N-(1-methoxypropan-2-yl)acrylamide (28)

87%).

Compound **28** was prepared according to the general procedure for *N*-SEM deprotection. *N*-[3-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]-*N*-(1-methoxypropan-2-yl)acrylamide (101 mg, 0.203 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **28** as a white solid (64.0 mg, 0.175 mmol,

Rf: 0.42 (5% MeOH in DCM); ¹H NMR (500 MHz, CDCl₃) δ 1.14 (d, *J* = 7.0 Hz, 3H, CH₃CHCH₂), 3.33 – 3.28 (m, 2H, CHCH₂OCH₃), 3.33 (s, 3H, CHCH₂OCH₃), 4.06 (s, 3H, ArOCH₃), 5.20 (br s, 1H, CH₃CHCH₂), 5.48 (d, *J* = 10.4 Hz, 1H, CCH=CH₂), 6.00 (dd, *J* = 16.7, 10.4 Hz, 1H, CCH=CH₂), 6.38 (d, *J* = 16.7 Hz, 1H, CCH=CH₂), 7.39 (s, 1H, ArH), 7.13 (s, 1H, ArH), 7.43 (dd [app. t], *J* = 7.7 Hz, 1H, ArH), 7.51 (s, 1H, ArH), 7.66 (d, *J* = 7.7 Hz, 1H, ArH), 8.52 (s, 1H, ArH), 12.37 (s, 1H, ArNH) ppm; ¹³C NMR (126 MHz, CDCl₃) δ 15.7, 49.8, 54.0, 58.6, 73.6, 103.2, 116.3, 121.9, 127.5, 128.4, 128.4, 129.1, 129.5, 131.3, 135.7, 137.9, 150.6, 153.1, 163.6, 166.0 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₀H₂₃N₄O₃: 367.1770; found: 367.1770.

N-(2-Ethoxyethyl)-N-[3-(4-methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]acrylamide (29)

Compound **29** was prepared according to the general procedure for *N*-SEM deprotection.

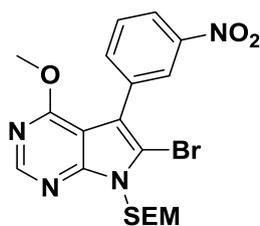
N-(2-Ethoxyethyl)-*N*-[3-(4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]acrylamide (154 mg, 0.310 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **29** as a white solid (92.0 mg, 0.251 mmol, 81%).

Rf: 0.41 (5% MeOH in DCM); ¹H NMR (500 MHz, CDCl₃) δ 1.10 (t, *J* = 7.0 Hz, 3H, OCH₂CH₃), 3.45 (q, *J* = 7.0 Hz, 2H, OCH₂CH₃), 3.66 (t, *J* = 6.2 Hz, 2H, NCH₂CH₂O), 4.01 (t, *J* = 6.2 Hz, 2H, NCH₂CH₂O), 4.06 (s, 3H, ArOCH₃), 5.53 (d, *J* = 10.4 Hz, 1H, CCH=CH₂), 6.18 (dd, *J* = 16.6, 10.4 Hz, 1H, CCH=CH₂), 6.40 (d, *J* = 16.8 Hz, 1H, CCH=CH₂), 7.14 (d, *J* = 7.8 Hz, 1H, ArH), 7.37 (s, 1H, ArH), 7.41 (dd [app. t], *J* = 7.8 Hz, 1H, ArH), 7.59 (s, 1H, ArH), 7.61 (d, *J* = 8.1 Hz, 1H, ArH), 8.51 (s, 1H, ArH), 12.42 (s, 1H, ArNH) ppm; ¹³C NMR (126 MHz, CDCl₃) δ 15.2, 49.3, 53.9, 66.2, 67.3, 103.2, 116.3, 121.7, 124.4, 126.3, 127.7, 127.8, 128.9, 129.2, 135.9, 142.1, 150.6, 153.1, 163.6, 165.8 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₀H₂₃N₄O₃: 367.1770; found: 367.1769.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

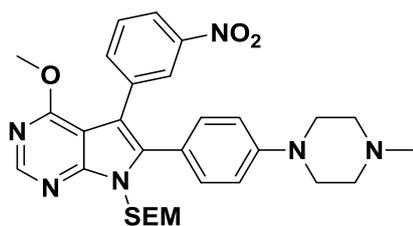
General procedure for pyrrolopyrimidine derivative bromination:²⁸

A flame-dried, 2-neck round-bottomed flask was charged with pyrrolopyrimidine derivative (1 equiv.), purged with Ar and dissolved in DCM (5 mL/mmol pyrrolopyrimidine derivative). A solution of *N*-bromosuccinimide (1.05 equiv.) in DCM (5 mL/mmol pyrrolopyrimidine derivative) was then added dropwise over a period of 15 min and the reaction flask was covered in foil. The reaction mixture was then allowed to stir at rt for 16 h or until complete consumption of the starting material as indicated by TLC, during which time an orange-brown colour developed. After addition of a sufficient amount of 10 mol% sodium thiosulphate so that the solution became clear, the reaction mixture was diluted with H₂O (20 mL/mmol pyrrolopyrimidine derivative) and DCM (20 mL/mmol pyrrolopyrimidine derivative) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 10 mL/mmol pyrrolopyrimidine derivative) and the combined organic layers were then washed with a saturated solution of brine (10 mL/mmol pyrrolopyrimidine derivative), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

6-Bromo-4-methoxy-5-(3-nitrophenyl)-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (30)

Compound **30** was prepared according to the general procedure for pyrrolopyrimidine derivative bromination. 4-Methoxy-5-(3-nitrophenyl)-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (2.00 g, 5.00 mmol) was brominated and purification was achieved using flash column chromatography with elution gradient of 0 – 30% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **30** as a white solid (2.20 g, 4.59 mmol, 92%).

R_f: 0.42 (15% EtOAc in PE); ¹H NMR (500 MHz, DMSO-*d*₆) δ –0.09 (s, 9H, CH₂Si(CH₃)₃), 0.86 (t, *J* = 7.6 Hz, 2H, CH₂CH₂Si), 3.59 (t, *J* = 7.6 Hz, 2H, OCH₂CH₂), 3.93 (s, 3H, ArOCH₃), 5.70 (s, 2H, NCH₂O), 7.77 (dd [app. t], *J* = 7.9 Hz, 1H, ArH), 7.98 (d, *J* = 6.8 Hz, 1H, ArH), 8.25 (d, *J* = 7.4 Hz, 1H, ArH), 8.35 (s, 1H, ArH), 8.54 (s, 1H, ArH) ppm; ¹³C NMR (500 MHz, DMSO-*d*₆) δ –1.4, 17.1, 54.0, 66.1, 71.9, 103.4, 112.9, 113.5, 122.3, 124.9, 129.5, 133.8, 137.0, 147.4, 151.7, 152.4, 161.6 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₉H₂₄BrN₄O₄Si: 479.0750; found: 479.0744.

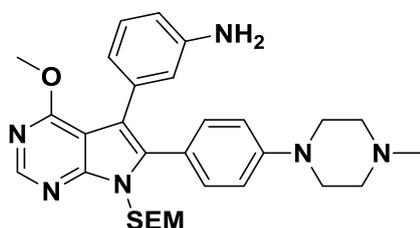
4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-5-(3-nitrophenyl)-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (31)

Compound **31** was prepared according to the general procedure for Suzuki-Miyaura cross coupling reaction. 6-Bromo-4-methoxy-5-(3-nitrophenyl)-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (400 mg, 0.834 mmol) and 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine (303 mg, 1.00 mmol) were coupled and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **31** as a yellow foam (444 mg, 0.772 mmol, 93%).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Rf: 0.32 (5% MeOH in DCM); ^1H NMR (600 MHz, $\text{DMSO-}d_6$) δ -0.12 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.79 (t, $J = 8.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 2.18 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.39 (m [app. t], $J = 4.6$ Hz, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.15 (m [app. t], $J = 4.6$ Hz, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.50 (t, $J = 8.2$ Hz, 2H, OCH_2CH_2), 3.93 (s, 3H, ArOCH_3), 5.45 (s, 2H, NCH_2O), 6.92 – 6.87 (m, 2H, $2 \times \text{ArH}$), 7.22 – 7.17 (m, 2H, $2 \times \text{ArH}$), 7.48 (dd [app. t], $J = 8.0$ Hz, 1H, ArH), 7.56 (d, $J = 7.8$ Hz, 1H, ArH), 8.04 – 8.02 (m, 1H, ArH), 8.12 (dd [app. t], $J = 1.9$ Hz, 1H, ArH), 8.50 (s, 1H, ArH) ppm; ^{13}C NMR (151 MHz, $\text{DMSO-}d_6$) δ -1.5, 17.3, 25.0, 45.7, 47.1, 53.7, 54.4, 66.0, 70.5, 73.5, 103.1, 110.8, 114.4, 118.0, 121.0, 125.3, 128.8, 128.9, 131.5, 131.8, 135.5, 137.1, 137.6, 147.2, 150.9, 150.9, 152.5, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{39}\text{N}_6\text{O}_4\text{Si}$: 575.2802; found: 575.2806.

3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (32)

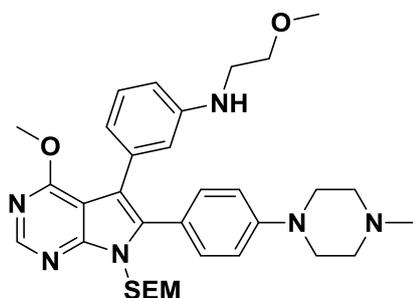


Compound **32** was prepared according to the general procedure for aromatic nitro derivative reduction. 4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-5-(3-nitrophenyl)-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (240 mg, 0.418 mmol) was reduced and purification was achieved using flash column chromatography with elution

gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **32** as a white foam (200 mg, 0.367 mmol, 88%).

Rf: 0.24 (5% MeOH in DCM); ^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ -0.11 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.79 (t, $J = 7.7$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 2.21 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.43 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.17 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.50 (t, $J = 7.7$ Hz, 2H, OCH_2CH_2), 3.91 (s, 3H, ArOCH_3), 4.92 (br s, 2H, ArNH_2), 5.45 (s, 2H, NCH_2O), 6.32 (d, $J = 7.0$ Hz, 1H, ArH), 6.40 (d, $J = 7.1$ Hz, 1H, ArH), 6.51 (s, 1H, ArH), 6.85 (d, $J = 7.9$ Hz, 1H, ArH), 6.91 – 6.87 (m, 2H, $2 \times \text{ArH}$), 7.24 – 7.17 (m, 2H, $2 \times \text{ArH}$), 8.46 (s, 1H, ArH) ppm; ^{13}C NMR (500 MHz, $\text{DMSO-}d_6$) δ -1.4, 17.3, 45.7, 47.1, 53.5, 54.5, 65.8, 70.5, 103.7, 112.3, 113.9, 114.2, 116.8, 118.9, 119.5, 127.8, 131.7, 134.4, 135.9, 147.8, 150.4, 150.4, 152.3, 162.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{41}\text{N}_6\text{O}_2\text{Si}$: 545.3060; found: 545.3061.

3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-N-(2-methoxyethyl)aniline (33)



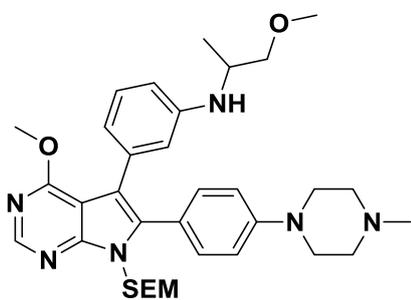
Compound **33** was prepared according to the general procedure for aniline derivative monoalkylation. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (114 mg, 0.209 mmol) was alkylated with 2-methoxyethyl 4-methylbenzenesulfonate (32.0 mg, 0.139 mmol) and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford

compound **33** as a white foam (45.0 mg, 74.6 μmol , 53% isolated, 81% brsm).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Rf: 0.39 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ -0.05 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.94 (t, J = 8.2 Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 2.42 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.66 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.16 (t, J = 5.1 Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.30 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.35 (s, 3H, CH_2OCH_3), 3.51 (t, J = 5.1 Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.64 (t, J = 8.2 Hz, 2H, OCH_2CH_2), 4.01 (s, 3H, ArOCH_3), 5.51 (s, 2H, NCH_2O), 6.49 (dd, J = 8.0, 1.9 Hz, 1H, ArH), 6.60 (s, 1H, ArH), 6.64 (d, J = 7.7 Hz, 1H, ArH), 6.86 – 6.83 (m, 2H, $2 \times \text{ArH}$), 7.03 (dd [app. t], J = 7.8 Hz, 1H, ArH), 7.31 – 7.29 (m, 2H, $2 \times \text{ArH}$), 8.49 (s, 1H, ArH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ -1.3, 18.1, 43.7, 45.9, 48.1, 53.7, 54.9, 58.9, 66.6, 70.9, 71.1, 104.5, 112.0, 114.5, 115.3, 115.9, 121.0, 121.4, 128.3, 132.4, 134.9, 136.5, 147.5, 150.5, 150.8, 153.0, 163.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{33}\text{H}_{47}\text{N}_6\text{O}_3\text{Si}$: 603.3479; found: 603.3480.

3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-N-(1-methoxypropan-2-yl)aniline (34)



Compound **34** was prepared according to the general procedure for aniline derivative monoalkylation. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (253 mg, 0.464 mmol) was alkylated with 1-methoxypropan-2-yl 4-methylbenzenesulfonate (76 mg, 0.311 mmol) and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford

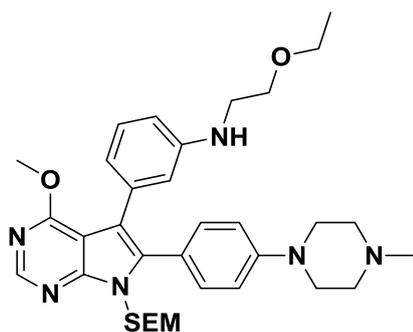
compound **34** as a white foam (111 mg, 0.180 mmol, 58% isolated, 86% brsm).

Rf: 0.43 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ -0.05 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.93 (t, J = 8.3 Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.81 – 1.73 (m, 3H, CH_3CHCH_2), 2.38 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.61 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.28 – 3.22 (m, 6H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$ and $\text{CHCH}_2\text{OCH}_3$), 3.31 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.44 (t, J = 8.3 Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.64 (m, 1H, CH_3CHCH_2), 4.01 (s, 3H, ArOCH_3), 5.51 (s, 2H, NCH_2O), 6.46 (dd, J = 8.0, 1.7 Hz, 1H, ArH), 6.55 (s, 1H, ArH), 6.59 (s, 1H, ArH), 6.60 (m, 1H, ArH), 6.86 – 6.83 (m, 2H, $2 \times \text{ArH}$), 7.01 (dd [app. t], J = 8.0 Hz, 1H, ArH), 7.31 – 7.28 (m, 2H, $2 \times \text{ArH}$), 8.49 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ -1.3, 18.1, 29.3, 41.8, 46.0, 48.0, 48.2, 53.7, 55.0, 58.7, 58.8, 66.6, 70.9, 71.2, 104.4, 111.5, 114.6, 115.2, 115.6, 120.5, 121.2, 128.3, 132.3, 134.8, 136.5, 147.7, 150.6, 150.7, 153.0, 163.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{34}\text{H}_{49}\text{N}_6\text{O}_3\text{Si}$: 617.3635; found: 617.3639.

N-(2-Ethoxyethyl)-3-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (35)

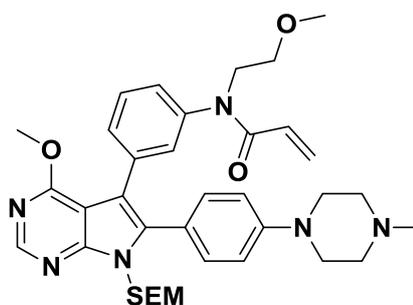
Compound **35** was prepared according to the general procedure for aniline derivative monoalkylation. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (267 mg, 0.490 mmol) was alkylated with 2-ethoxyethyl 4-methylbenzenesulfonate (80.0 mg, 0.327 mmol) and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **35** as a white foam (121 mg, 0.196 mmol, 59% isolated, 85% brsm).

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



Rf: 0.41 (5% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ -0.05 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.94 (t, $J = 8.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.20 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.39 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.63 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.17 (t, $J = 5.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.28 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.49 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.55 (t, $J = 5.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.64 (t, $J = 8.2$ Hz, 2H, OCH_2CH_2), 4.01 (s, 3H, ArOCH_3), 5.51 (s, 2H, NCH_2O), 6.50 (dd, $J = 8.0, 1.8$ Hz, 1H, ArH), 6.62 (s, 1H, ArH), 6.63 (d, $J = 7.9$ Hz, 1H, ArH), 6.86 – 6.83 (m, 2H, $2 \times \text{ArH}$), 7.02 (dd [app. t], $J = 7.7$ Hz, 1H, ArH), 7.32 – 7.28 (m, 2H, $2 \times \text{ArH}$), 8.49 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ -1.3, 15.3, 18.1, 43.8, 46.0, 48.2, 53.7, 55.0, 66.5, 66.6, 69.0, 70.9, 104.5, 112.0, 114.5, 115.2, 116.0, 121.0, 121.3, 128.3, 132.4, 134.9, 136.5, 147.6, 150.6, 150.7, 153.0, 163.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{34}\text{H}_{49}\text{N}_6\text{O}_3\text{Si}$: 617.3635; found: 617.3638.

N-(3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)-N-(2-methoxyethyl)acrylamide (36)



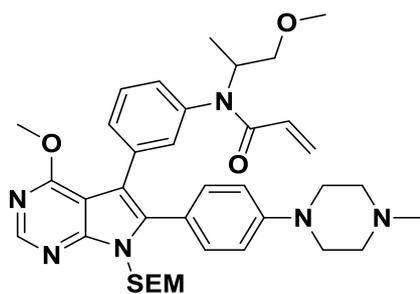
Compound **36** was prepared according to the general procedure for aniline derivative acylation. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-N-(2-methoxyethyl)aniline (73.0 mg, 0.121 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **36** as a white solid (51.0 mg, 77.6 μmol , 65%).

Rf: 0.33 (5% MeOH in DCM); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ -0.03 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.95 (t, $J = 8.4$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 2.44 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.68 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.30 (s, 3H, CH_2OCH_3), 3.32 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.49 (t, $J = 6.1$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.67 (t, $J = 8.4$ Hz, 2H, OCH_2CH_2), 3.86 (t, $J = 6.1$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 4.02 (s, 3H, ArOCH_3), 5.48 (dd, $J = 10.3, 2.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 5.52 (s, 2H, NCH_2O), 6.01 (dd, $J = 16.8, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.33 (dd, $J = 16.8, 2.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.88 – 6.84 (m, 2H, $2 \times \text{ArH}$), 7.05 – 7.01 (m, 1H, ArH), 7.13 – 7.11 (m, 1H, ArH), 7.26 (s, 1H, ArH), 7.30 – 7.28 (m, 2H, $2 \times \text{ArH}$), 7.31 (dd, $J = 6.1, 1.7$ Hz, 1H, ArH), 8.52 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ -1.3, 18.1, 45.9, 47.9, 48.7, 53.8, 54.9, 58.7, 66.8, 69.3, 71.0, 104.2, 113.0, 115.4, 120.5, 126.2, 127.6, 128.6, 128.8, 130.7, 131.0, 132.4, 135.9, 137.2, 141.3, 150.8, 151.1, 153.1, 163.1, 165.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{36}\text{H}_{49}\text{N}_6\text{O}_4\text{Si}$: 657.3585; found: 657.3591.

N-(3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)-N-(1-methoxypropan-2-yl)acrylamide (37)

Compound **37** was prepared according to the general procedure for aniline derivative acylation. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-N-(1-methoxypropan-2-yl)aniline (111 mg, 0.180 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM.

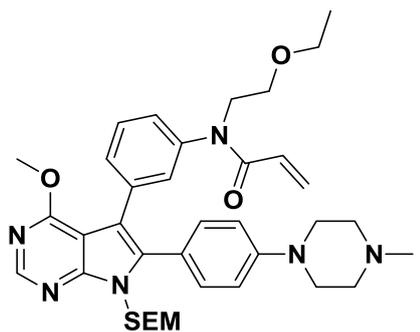
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



Pure fractions were evaporated to dryness to afford compound **37** as a white solid (81.0 mg, 0.121 mmol, 67%).

Rf: 0.36 (5% MeOH in DCM); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ -0.03 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.95 (t, $J = 8.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.81 – 1.74 (m, 3H, CH_3CHCH_2), 2.45 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.70 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.27 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.34 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.42 – 3.35 (m, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.67 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.80 – 3.74 (m, 1H, CH_3CHCH_2), 4.02 (s, 3H, ArOCH_3), 5.46 (dd, $J = 10.3, 1.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 5.52 (s, 2H, NCH_2O), 5.98 (dd, $J = 16.8, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.32 (dd, $J = 16.8, 1.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.88 – 6.85 (m, 2H, $2 \times \text{ArH}$), 7.00 – 6.96 (m, 1H, ArH), 7.09 (s, 1H, ArH), 7.26 (s, 1H, ArH), 7.29 (m, 2H, $2 \times \text{ArH}$), 7.31 (dd, $J = 7.0, 0.8$ Hz, 1H, ArH), 8.52 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ -1.3, 18.1, 28.1, 45.8, 47.0, 47.8, 53.8, 54.8, 58.7, 66.8, 70.5, 71.0, 104.2, 112.9, 115.4, 120.6, 126.0, 127.4, 128.7, 128.9, 130.7, 131.0, 132.4, 135.9, 137.2, 141.1, 150.7, 151.1, 153.1, 163.1, 165.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{37}\text{H}_{51}\text{N}_6\text{O}_4\text{Si}$: 671.3741; found: 671.3741.

***N*-(2-Ethoxyethyl)-*N*-(3-[4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[2-(trimethylsilyl)ethoxy]methyl]-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl]phenyl)acrylamide (**38**)**

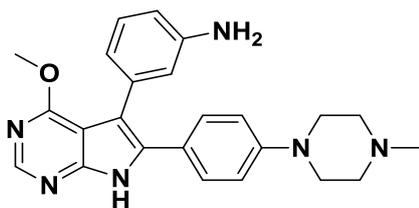


Compound **38** was prepared according to the general procedure for aniline derivative acylation. *N*-(2-Ethoxyethyl)-3-[4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[2-(trimethylsilyl)ethoxy]methyl]-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl]aniline (165 mg, 0.267 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **38** as a white solid (160 mg, 0.238 mmol, 89%).

89%).

Rf: 0.36 (5% MeOH in DCM); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ -0.02 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.96 (t, $J = 8.4$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.13 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.52 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.80 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.40 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.45 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.57 (t, $J = 6.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.68 (t, $J = 8.4$ Hz, 2H, OCH_2CH_2), 3.88 (t, $J = 6.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 4.03 (s, 3H, ArOCH_3), 5.48 (dd, $J = 10.3, 2.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 5.53 (s, 2H, NCH_2O), 6.01 (dd, $J = 16.7, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.34 (dd, $J = 16.8, 2.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.89 – 6.86 (m, 2H, $2 \times \text{ArH}$), 7.08 – 7.04 (m, 1H, ArH), 7.16 – 7.14 (m, 1H, ArH), 7.28 (s, 1H, ArH), 7.30 – 7.28 (m, 2H, $2 \times \text{ArH}$), 7.30 (dd, $J = 7.0, 0.8$ Hz, 1H, ArH), 8.54 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ -1.3, 15.3, 18.1, 45.4, 47.6, 49.1, 53.8, 54.6, 66.3, 66.8, 67.3, 71.0, 104.2, 113.1, 115.6, 120.9, 126.3, 127.5, 128.6, 128.9, 130.6, 131.0, 132.5, 135.8, 137.0, 141.5, 150.6, 151.1, 153.1, 163.1, 165.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{37}\text{H}_{51}\text{N}_6\text{O}_4\text{Si}$: 671.3741; found: 671.3746.

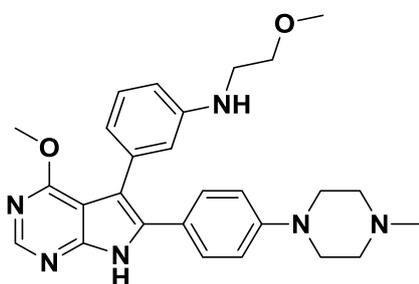
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (39)

Compound **39** was prepared according to the general procedure for *N*-SEM deprotection. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (180 mg, 0.330 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in

DCM. Pure fractions were evaporated to dryness to afford compound **39** as a white solid (111 mg, 0.268 mmol, 81%).

Rf: 0.43 (10% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 2.34 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.53 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.10 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.69 (br s, 2H, ArNH_2), 3.78 (s, 3H, ArOCH_3), 6.58 (s, 1H, ArH), 6.61 (d, $J = 8.1$ Hz, 1H, ArH), 6.69 (s, 1H, ArH), 6.76 – 6.70 (m, 2H, $2 \times \text{ArH}$), 6.86 (d, $J = 7.2$ Hz, 1H, ArH), 7.13 – 7.09 (m, 2H, $2 \times \text{ArH}$), 7.13 (d, $J = 7.5$ Hz, 1H, ArH), 8.18 (s, 1H, ArH), 13.28 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 46.1, 47.9, 53.5, 55.1, 106.0, 111.2, 113.3, 115.2, 118.4, 121.9, 128.7, 129.1, 134.0, 135.9, 146.1, 149.6, 150.2, 152.1, 163.0, 171.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{24}\text{H}_{27}\text{N}_6\text{O}$: 415.2246; found: 415.2245.

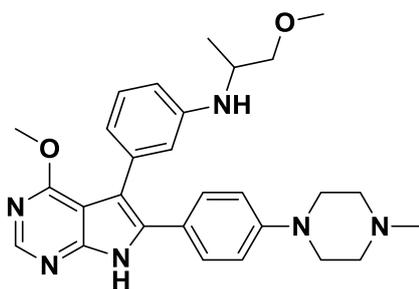
3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-*N*-(2-methoxyethyl)aniline (40)

Compound **40** was prepared according to the general procedure for *N*-SEM deprotection. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-*N*-(2-methoxyethyl)aniline (60.0 mg, 0.100 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to

dryness to afford compound **40** as a white solid (36.0 mg, 76.1 μmol , 76%).

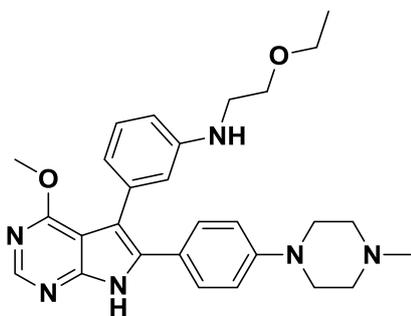
Rf: 0.48 (10% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 2.40 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.62 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.24 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.29 (t, $J = 6.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.37 (s, 3H, CH_2OCH_3), 3.58 (t, $J = 6.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.93 (s, 3H, ArOCH_3), 6.61 (d, $J = 7.6$ Hz, 1H, ArH), 6.79 (s, 1H, ArH), 6.85 – 6.81 (m, 2H, $2 \times \text{ArH}$), 7.16 (dd [app. t], $J = 7.5$ Hz, 1H, ArH), 7.34 – 7.30 (m, 2H, $2 \times \text{ArH}$), 8.34 (s, 1H, ArH), 13.15 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 43.8, 46.0, 48.0, 53.7, 55.0, 58.9, 71.2, 106.2, 111.9, 111.9, 115.4, 116.1, 121.1, 122.8, 128.7, 129.4, 133.9, 135.6, 147.9, 149.8, 150.3, 152.3, 163.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{33}\text{N}_6\text{O}_2$: 473.2665; found: 473.2661.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-N-(1-methoxypropan-2-yl)aniline (41)

Compound **41** was prepared according to the general procedure for *N*-SEM deprotection. 3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-*N*-(1-methoxypropan-2-yl)aniline (68.0 mg, 0.110 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **41** as a white solid (41.0 mg, 84.2 μ mol, 77%).

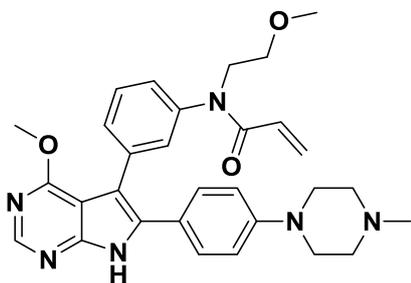
Rf: 0.54 (10% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 1.18 (d, $J = 6.4$ Hz, 3H, CH_3CHCH_2), 2.44 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.67 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.28 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.34 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.42 – 3.35 (m, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.67 – 3.60 (m, 1H, CH_3CHCH_2), 3.95 (s, 3H, ArOCH_3), 6.59 (d, $J = 8.1$ Hz, 1H, ArH), 6.76 (s, 1H, ArH), 6.79 (d, $J = 7.6$ Hz, 1H, ArH), 6.88 – 6.82 (m, 2H, $2 \times \text{ArH}$), 7.14 (dd [app. t], $J = 7.6$ Hz, 1H, ArH), 7.37 – 7.32 (m, 2H, $2 \times \text{ArH}$), 8.35 (s, 1H, ArH), 13.05 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 18.0, 45.9, 48.0, 48.5, 53.7, 54.9, 59.2, 76.2, 106.1, 112.0, 112.7, 115.6, 116.3, 120.8, 123.1, 128.8, 129.5, 133.9, 135.6, 147.2, 149.9, 150.3, 152.3, 163.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{35}\text{N}_6\text{O}_2$: 487.2821; found: 487.2821.

N-(2-Ethoxyethyl)-3-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (42)

Compound **42** was prepared according to the general procedure for *N*-SEM deprotection. *N*-(2-Ethoxyethyl)-3-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (74.0 mg, 0.120 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **42** as a white solid (44.0 mg, 90.4 μ mol, 76%).

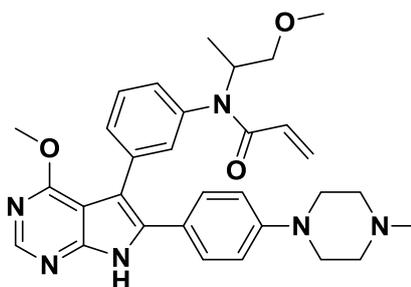
Rf: 0.55 (10% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 1.21 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.41 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.64 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.26 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.28 (t, $J = 5.1$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.52 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.62 (t, $J = 5.1$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.93 (s, 3H, ArOCH_3), 6.79 (s, 1H, ArH), 6.81 (d, $J = 7.8$ Hz, 1H, ArH), 6.88 – 6.82 (m, 2H, $2 \times \text{ArH}$), 7.16 (dd [app. t], $J = 7.8$ Hz, 1H, ArH), 7.35 – 7.31 (m, 2H, $2 \times \text{ArH}$), 8.35 (s, 1H, ArH), 13.14 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 15.3, 43.9, 46.0, 48.0, 53.7, 55.0, 66.5, 69.1, 106.2, 111.9, 111.9, 115.4, 116.1, 121.0, 122.9, 128.7, 129.4, 133.9, 135.6, 148.0, 149.8, 150.3, 152.3, 163.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{35}\text{N}_6\text{O}_2$: 487.2821; found: 487.2820.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-(3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)-N-(2-methoxyethyl)acrylamide (43)

Compound **43** was prepared according to the general procedure for *N*-SEM deprotection. *N*-(3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)-*N*-(2-methoxyethyl)acrylamide (51.0 mg, 77.6 μ mol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **43** as a white solid (32.0 mg, 60.7 μ mol, 79%).

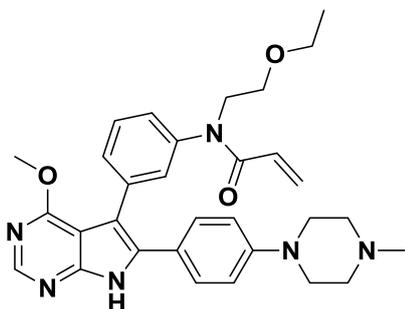
Rf: 0.40 (10% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 2.40 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.64 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.29 (s, 3H, CH_2OCH_3), 3.30 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.56 (t, $J = 6.0$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.95 (t, $J = 6.0$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.99 (s, 3H, ArOCH_3), 5.50 (dd, $J = 10.3, 1.6$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.14 (dd, $J = 16.7, 10.3$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.36 (dd, $J = 16.8, 1.6$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.91 – 6.88 (m, 2H, $2 \times \text{ArH}$), 7.12 (d, $J = 7.8$ Hz, 1H, ArH), 7.30 (s, 1H, ArH), 7.34 (dd [app. t], $J = 7.7$ Hz, 1H, ArH), 7.37 – 7.35 (m, 2H, $2 \times \text{ArH}$), 7.40 (d, $J = 7.7$ Hz, 1H, ArH), 8.37 (s, 1H, ArH), 13.28 (s, 1H, ArNH), ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 46.0, 48.1, 48.9, 53.8, 54.9, 58.7, 69.5, 105.8, 110.3, 115.6, 122.2, 126.5, 127.6, 128.9, 129.0, 129.6, 130.8, 131.1, 134.7, 136.4, 141.7, 150.1, 150.8, 152.5, 163.2, 165.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{35}\text{N}_6\text{O}_3$: 527.2771; found: 527.2773.

N-(3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)-N-(1-methoxypropan-2-yl)acrylamide (44)

Compound **44** was prepared according to the general procedure for *N*-SEM deprotection. *N*-(3-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)-*N*-(1-methoxypropan-2-yl)acrylamide (47.0 mg, 70.1 μ mol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **44** as a white solid (28.0 mg, 51.8 μ mol, 75%).

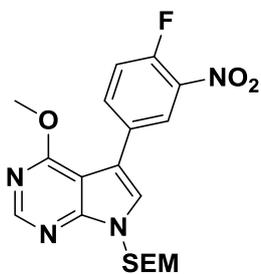
Rf: 0.42 (10% MeOH in DCM); ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 1.68 – 1.60 (m, 3H, CH_3CHCH_2), 2.22 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.44 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.15 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.16 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.30 – 3.26 (m, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.73 – 3.68 (m, 1H, CH_3CHCH_2), 3.89 (s, 3H, ArOCH_3), 5.56 (d, $J = 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.01 (dd, $J = 16.6, 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.13 (dd, $J = 16.6, 2.1$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.90 – 6.86 (m, 2H, $2 \times \text{ArH}$), 7.11 (s, 1H, ArH), 7.16 (d, $J = 7.3$ Hz, 1H, ArH), 7.21 – 7.19 (m, 2H, $2 \times \text{ArH}$), 7.33 (d, $J = 7.7$ Hz, 1H, ArH), 7.42 (dd [app. t], $J = 7.7$ Hz, 1H, ArH), 8.38 (s, 1H, ArH), 12.43 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 27.6, 45.7, 46.2, 47.2, 53.4, 54.5, 57.9, 69.6, 104.5, 109.2, 114.6, 120.6, 126.1, 127.4, 128.9, 129.2, 129.4, 130.3, 130.8, 134.0, 136.3, 141.1, 150.3, 150.3, 152.3, 162.2, 164.0 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{31}\text{H}_{37}\text{N}_6\text{O}_3$: 541.2927; found: 541.2928.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-(2-Ethoxyethyl)-N-(3-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)acrylamide (45)

Compound **45** was prepared according to the general procedure for *N*-SEM deprotection. *N*-(2-Ethoxyethyl)-*N*-(3-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7*H*-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl)acrylamide (60.0 mg, 89.4 μ mol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **45** as a white solid (36.0 mg, 66.6 μ mol, 73%).

Rf: 0.41 (10% MeOH in DCM); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 1.11 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.44 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.70 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.34 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.44 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.62 (t, $J = 6.1$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 3.95 (t, $J = 6.1$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{O}$), 4.00 (s, 3H, ArOCH_3), 5.50 (d, $J = 10.6$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.15 (dd, $J = 16.8, 10.6$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.36 (dd, $J = 16.8, 1.9$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.92 – 6.89 (m, 2H, $2 \times \text{ArH}$), 7.13 (d, $J = 7.8$ Hz, 1H, ArH), 7.31 (s, 1H, ArH), 7.34 (dd [app. t], $J = 7.7$ Hz, 1H, ArH), 7.38 – 7.36 (m, 2H, $2 \times \text{ArH}$), 7.39 (d, $J = 7.7$ Hz, 1H, ArH), 8.39 (s, 1H, ArH), 13.03 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 15.3, 45.9, 48.0, 49.3, 53.8, 54.9, 66.3, 67.4, 105.8, 110.4, 115.7, 122.4, 126.5, 127.6, 128.9, 129.0, 129.7, 130.7, 131.2, 134.5, 136.3, 142.0, 150.2, 150.7, 152.5, 163.3, 165.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{31}\text{H}_{37}\text{N}_6\text{O}_3$: 541.2927; found: 541.2929.

5-(4-Fluoro-3-nitrophenyl)-4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (46)

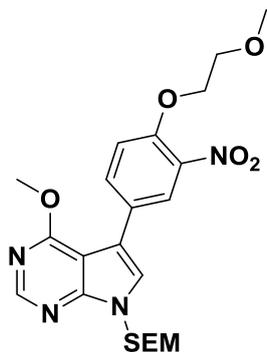
Compound **46** was prepared according to the general procedure for Suzuki-Miyaura cross coupling reaction. 5-Iodo-4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7*H*-pyrrolo[2,3-d]pyrimidine (1.00 g, 2.47 mmol) and (4-fluoro-3-nitrophenyl)boronic acid (548 mg, 2.96 mmol) were coupled and purification was achieved using flash column chromatography with elution gradient of 0 – 50% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **46** as a yellow semi-solid (890 mg, 2.12 mmol, 86%).

Rf: 0.63 (30% EtOAc in PE); $^1\text{H NMR}$ (700 MHz, $\text{DMSO}-d_6$) δ -0.09 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.84 (t, $J = 8.0$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.56 (t, $J = 8.0$, 2H, OCH_2CH_2), 4.05 (s, 3H, ArOCH_3), 5.61 (s, 2H, NCH_2O), 7.62 (dd, $J = 11.1, 8.8$ Hz, 1H, ArH), 8.03 (s, 1H, ArH), 8.11 – 8.07 (m, 1H, ArH), 8.49 (dd, $J = 7.3, 2.2$ Hz, 1H, ArH), 8.50 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (176 MHz, $\text{DMSO}-d_6$) δ -1.5, 17.1, 53.8, 65.9, 72.8, 102.1, 112.8, 118.4 (d, $J = 20.8$ Hz), 125.2 (d, $J = 2.3$ Hz), 126.7, 130.9 (d, $J = 3.8$ Hz), 135.5 (d, $J = 8.4$ Hz), 136.6 (d, $J = 7.7$ Hz), 151.3, 152.5, 152.7, 162.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{24}\text{N}_4\text{O}_4\text{SiF}$: 419.1551; found: 419.1550.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

General procedure for nucleophilic aromatic substitution:²⁹

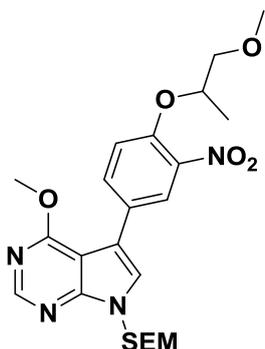
A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in oil, 1.7 equiv.), purged with Ar and to this was added THF (10 mL/mmol fluorinated aromatic derivative). The resulting suspension was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) and to this was added a solution of glycol ether (1.2 equiv.) in THF (5 mL/mmol fluorinated aromatic derivative) dropwise. The solution was then allowed to warm to rt and stirred for 1 h. After recooling the solution to $-10\text{ }^{\circ}\text{C}$ (acetone/ice), a solution of fluorinated aromatic derivative (1 equiv.) in THF (5 mL/mmol fluorinated aromatic derivative) was added dropwise and the reaction mixture was then allowed to warm to rt and stirred for 16 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (20 mL/mmol fluorinated aromatic derivative) and EtOAc (20 mL/mmol fluorinated aromatic derivative) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc ($3 \times 10\text{ mL/mmol}$ fluorinated aromatic derivative) and the combined organic layers were then washed with a saturated solution of brine ($3 \times 20\text{ mL/mmol}$ fluorinated aromatic derivative), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

4-Methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (47)

Compound **47** was prepared according to the general procedure for general procedure for nucleophilic aromatic substitution. 5-(4-Fluoro-3-nitrophenyl)-4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (440 mg, 1.05 mmol) was reacted with 2-methoxyethanol (95.0 mg, 1.26 mmol) and purification was achieved using flash column chromatography with elution gradient of 0 – 50% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **46** as a yellow semi-solid (354 mg, 0.746 mmol, 71%).

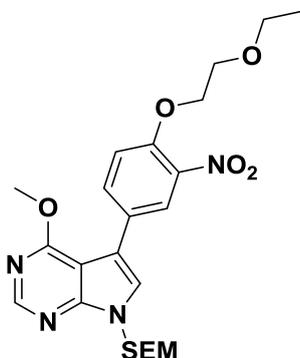
Rf: 0.29 (20% EtOAc in PE); ^1H NMR (500 MHz, CDCl_3) δ -0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.92 (t, $J = 8.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.46 (s, 3H, CH_2OCH_3), 3.56 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.82 (t, $J = 5.0$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.10 (s, 3H, ArOCH_3), 4.29 (t, $J = 5.0$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.65 (s, 2H, NCH_2O), 7.14 (d, $J = 8.7$ Hz, 1H, ArH), 7.30 (s, 1H, ArH), 7.78 (dd, $J = 8.7, 2.2$ Hz, 1H, ArH), 8.23 (d, $J = 2.2$ Hz, 1H, ArH), 8.51 (s, 1H, ArH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ -1.3 , 17.8, 54.1, 59.6, 66.8, 69.8, 70.8, 73.2, 103.2, 115.1, 115.5, 123.7, 126.0, 127.1, 134.0, 134.0, 151.2, 151.6, 153.0, 163.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{22}\text{H}_{31}\text{N}_4\text{O}_6\text{Si}$: 475.2013; found: 475.2013.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

4-Methoxy-5-{4-[(1-methoxypropan-2-yl)oxy]-3-nitrophenyl}-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (48)

Compound **48** was prepared according to the general procedure for general procedure for nucleophilic aromatic substitution. 5-(4-Fluoro-3-nitrophenyl)-4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (500 mg, 1.19 mmol) was reacted with 1-methoxypropan-2-ol (129 mg, 1.43 mmol) and purification was achieved using flash column chromatography with elution gradient of 0 – 50% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **48** as a yellow semi-solid (425 mg, 0.870 mmol, 73%).

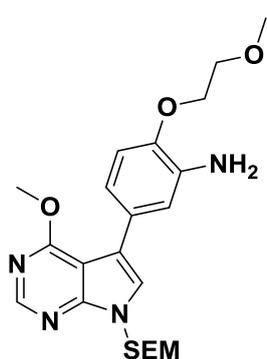
Rf: 0.37 (20% EtOAc in PE); ^1H NMR (700 MHz, CDCl_3) δ -0.05 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.93 (t, $J = 8.4$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.40 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 3.42 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.56 – 3.54 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.57 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.68 – 3.64 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 4.12 (s, 3H, ArOCH_3), 4.71 – 4.68 (m, 1H, CH_3CHCH_2), 5.66 (s, 2H, NCH_2O), 7.20 (d, $J = 8.7$ Hz, 1H, ArH), 7.31 (s, 1H, ArH), 7.76 (dd, $J = 8.7, 2.1$ Hz, 1H, ArH), 8.18 (d, $J = 2.1$ Hz, 1H, ArH), 8.52 (s, 1H, ArH) ppm; ^{13}C NMR (176 MHz, CDCl_3) δ -1.3, 17.1, 17.9, 54.1, 59.6, 66.8, 73.3, 76.1, 76.5, 103.3, 115.7, 116.8, 123.7, 125.8, 127.0, 133.7, 141.0, 150.5, 151.5, 152.9, 163.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{33}\text{N}_4\text{O}_6\text{Si}$: 489.2169; found: 489.2170.

5-[4-(2-Ethoxyethoxy)-3-nitrophenyl]-4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (49)

Compound **49** was prepared according to the general procedure for general procedure for nucleophilic aromatic substitution. 5-(4-Fluoro-3-nitrophenyl)-4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (500 mg, 1.19 mmol) was alkylated with 2-ethoxyethanol (129 mg, 1.43 mmol) and purification was achieved using flash column chromatography with elution gradient of 0 – 50% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **49** as a yellow semi-solid (430 mg, 0.881 mmol, 74%).

Rf: 0.38 (20% EtOAc in PE); ^1H NMR (700 MHz, CDCl_3) δ -0.04 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.94 (t, $J = 8.4$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.24 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 3.58 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.64 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.87 (t, $J = 4.6$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.13 (s, 3H, ArOCH_3), 4.31 (t, $J = 4.6$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.67 (s, 2H, NCH_2O), 7.16 (d, $J = 8.7$ Hz, 1H, ArH), 7.32 (s, 1H, ArH), 7.78 (dd, $J = 8.6, 2.2$ Hz, 1H, ArH), 8.23 (d, $J = 2.2$ Hz, 1H, ArH), 8.53 (s, 1H, ArH) ppm; ^{13}C NMR (176 MHz, CDCl_3) δ -1.3, 15.3, 17.9, 54.3, 66.9, 67.3, 68.8, 70.0, 73.4, 103.3, 115.2, 115.7, 123.8, 126.0, 126.9, 134.0, 140.0, 151.3, 151.4, 152.7, 163.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{33}\text{N}_4\text{O}_6\text{Si}$: 489.2169; found: 489.2172.

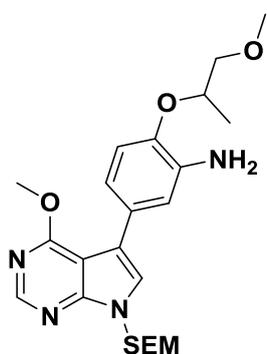
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

5-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-(2-methoxyethoxy)aniline (50)

Compound **50** was prepared according to the general procedure for aromatic nitro derivative reduction. 4-Methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (125 mg, 0.263 mmol) was reduced and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **50** as a yellow solid (112 mg, 0.252 mmol, 96%).

Rf: 0.52 (2% MeOH in DCM); ¹H NMR (400 MHz, CDCl₃) δ -0.06 (s, 9H, CH₂Si(CH₃)₃), 0.91 (t, *J* = 8.4 Hz, 2H, CH₂CH₂Si), 3.43 (s, 3H, CH₂OCH₃), 3.55 (t, *J* = 8.4 Hz, 2H,

OCH₂CH₂Si), 3.78 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 4.08 (s, 3H, ArOCH₃), 4.20 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 4.45 (br s, 2H, ArNH₂), 5.62 (s, 2H, NCH₂O), 6.88 (d, *J* = 8.4 Hz, 1H, ArH), 7.10 (dd, *J* = 8.3, 2.1 Hz, 1H, ArH), 7.20 (s, 1H, ArH), 7.21 (d, *J* = 2.1 Hz, 1H, ArH), 8.47 (s, 1H, ArH) ppm; ¹³C NMR (101 MHz, CDCl₃) δ -1.3, 17.8, 53.5, 59.2, 66.6, 68.5, 71.2, 73.1, 103.5, 112.8, 117.7, 117.7, 121.1, 123.3, 127.7, 133.7, 146.6, 151.3, 153.0, 163.6 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₂H₃₃N₄O₄Si: 445.2271; found: 445.2266.

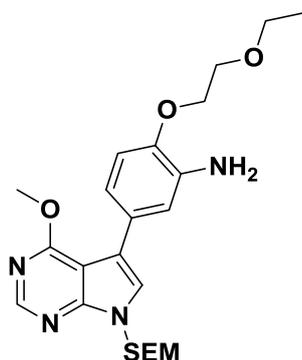
5-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]aniline (51)

Compound **51** was prepared according to the general procedure for aromatic nitro derivative reduction. 4-Methoxy-5-[4-[(1-methoxypropan-2-yl)oxy]-3-nitrophenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (102 mg, 0.209 mmol) was reduced and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **51** as a yellow solid (89.0 mg, 0.194 mmol, 92%).

Rf: 0.49 (2% MeOH in DCM); ¹H NMR (400 MHz, CDCl₃) δ 0.02 (s, 9H, CH₂Si(CH₃)₃),

0.99 (t, *J* = 8.2 Hz, 2H, CH₂CH₂Si), 1.43 (d, *J* = 6.3 Hz, 3H, CH₃CHCH₂), 3.50 (s, 3H, CHCH₂OCH₃), 3.60 – 3.55 (m, 1H, CHCH₂OCH₃), 3.63 (t, *J* = 8.2 Hz, 2H, OCH₂CH₂Si), 3.72 – 3.66 (m, 1H, CHCH₂OCH₃), 4.16 (s, 3H, ArOCH₃), 4.20 (br s, 2H, ArNH₂), 4.53 (m, 1H, CH₃CHCH₂), 5.69 (s, 2H, NCH₂O), 6.99 (d, *J* = 8.3 Hz, 1H, ArH), 7.08 (dd, *J* = 8.3, 2.1 Hz, 1H, ArH), 7.15 (d, *J* = 2.1 Hz, 1H, ArH), 7.27 (s, 1H, ArH), 8.55 (s, 1H, ArH); ¹³C NMR (101 MHz, CDCl₃) δ -1.4, 17.3, 17.8, 53.7, 59.3, 66.5, 73.1, 75.0, 76.1, 103.4, 115.8, 116.5, 117.9, 119.3, 123.1, 128.0, 137.3, 144.8, 151.2, 153.0, 163.5 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₃H₃₅N₄O₄Si: 459.2428; found: 459.2428.

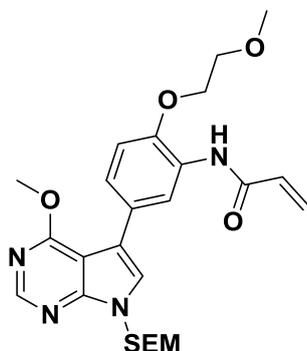
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

2-(2-Ethoxyethoxy)-5-(4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline**(52)**

Compound **52** was prepared according to the general procedure for aromatic nitro derivative reduction. 5-[4-(2-Ethoxyethoxy)-3-nitrophenyl]-4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (118 mg, 0.241 mmol) was reduced and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **52** as a yellow solid (98.0 mg, 0.214 mmol, 89%).

Rf: 0.46 (2% MeOH in DCM); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ -0.07 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.90 (t, $J = 8.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.23 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3),

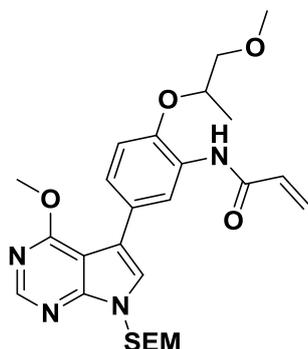
3.54 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.59 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.80 (t, $J = 5.0$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.06 (s, 3H, ArOCH_3), 4.18 (t, $J = 5.0$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.50 (br s, 2H, ArNH_2), 5.60 (s, 2H, NCH_2O), 6.85 (d, $J = 8.3$ Hz, 1H, ArH), 7.03 (dd, $J = 8.3, 2.1$ Hz, 1H, ArH), 7.09 – 7.07 (m, 1H, ArH), 7.17 (s, 1H, ArH), 8.47 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ -1.4, 15.3, 17.8, 53.7, 66.5, 66.8, 68.6, 69.1, 73.0, 103.4, 112.7, 116.7, 117.8, 119.8, 123.1, 127.5, 135.5, 146.1, 151.2, 152.9, 163.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{35}\text{N}_4\text{O}_4\text{Si}$: 459.2428; found: 459.2430.

N-[5-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-(2-methoxyethoxy)phenyl]acrylamide (53)

Compound **53** was prepared according to the general procedure for aniline derivative acylation. 5-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-(2-methoxyethoxy)aniline (49.0 mg, 0.110 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **53** as a white solid (49.0 mg, 98.2 μmol , 89%).

Rf: 0.65 (5% MeOH in DCM); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ -0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.92 (t, $J = 8.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.46 (s, 3H, CH_2OCH_3), 3.56 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.74 (t, $J = 5.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.13 (s, 3H, ArOCH_3), 4.22 (t, $J = 5.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.64 (s, 2H, NCH_2O), 5.75 (dd, $J = 10.2, 1.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.31 (dd, $J = 16.8, 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.42 (d, $J = 16.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.00 (d, $J = 8.4$ Hz, 1H, ArH), 7.32 (s, 1H, ArH), 7.38 (dd, $J = 8.4, 2.1$ Hz, 1H, ArH), 8.32 (s, 1H, ArH), 8.50 (s, 1H, ArH), 8.86 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ -1.3, 17.8, 54.1, 59.2, 66.6, 70.1, 71.0, 73.3, 103.5, 114.2, 117.8, 120.6, 123.9, 124.4, 127.2, 128.4, 129.3, 131.9, 146.4, 151.0, 152.7, 163.5, 163.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{25}\text{H}_{35}\text{N}_4\text{O}_5\text{Si}$: 499.2377; found: 499.2384.

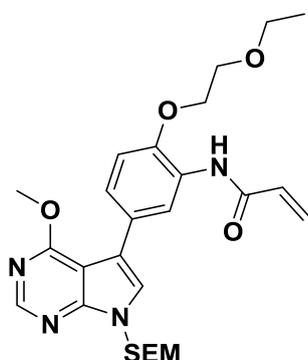
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-[5-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]phenyl]acrylamide (54)

Compound **54** was prepared according to the general procedure for aniline derivative acylation.

5-(4-Methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]aniline (55.0 mg, 0.120 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **54** as a white solid (54.0 mg, 0.105 mmol, 88%).

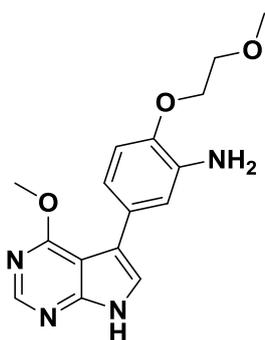
Rf: 0.66 (5% MeOH in DCM); ¹H NMR (600 MHz, CDCl₃) δ 0.00 (s, 9H, CH₂Si(CH₃)₃), 0.98 (t, *J* = 8.2 Hz, 2H, CH₂CH₂Si), 1.44 (d, *J* = 6.4 Hz, 3H, CH₃CHCH₂), 3.52 (s, 3H, CHCH₂OCH₃), 3.59 – 3.56 (m, 1H, CHCH₂OCH₃), 3.61 (t, *J* = 8.2 Hz, 2H, OCH₂CH₂Si), 3.68 – 3.63 (m, 1H, CHCH₂OCH₃), 4.20 (s, 3H, ArOCH₃), 4.38 – 4.34 (m, 1H, CH₃CHCH₂), 5.70 (s, 2H, NCH₂O), 5.80 (dd, 1H, *J* = 10.2, 1.0 Hz, 1H, CCH=CH₂), 6.36 (dd, *J* = 16.8, 10.2 Hz, 1H, CCH=CH₂), 6.49 (d, *J* = 16.8 Hz, 1H, CCH=CH₂), 7.11 (d, *J* = 8.4 Hz, 1H, ArH), 7.39 (s, 1H, ArH), 7.42 (dd, *J* = 8.4, 2.1 Hz, 1H, ArH), 8.56 (s, 1H, ArH), 8.76 (s, 1H, ArH), 8.96 (s, 1H, ArNH), ppm; ¹³C NMR (151 MHz, CDCl₃) δ –1.3, 17.1, 17.8, 54.1, 59.4, 66.6, 73.3, 76.0, 77.7, 103.5, 117.8, 118.2, 120.7, 124.0, 124.3, 127.2, 129.2, 130.8, 131.8, 145.7, 150.9, 152.7, 163.6, 163.6 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₆H₃₇N₄O₅Si: 513.2533; found: 513.2536.

N-[2-(2-Ethoxyethoxy)-5-(4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]acrylamide (55)

Compound **55** was prepared according to the general procedure for aniline derivative acylation. 2-(2-Ethoxyethoxy)-5-(4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (56.0 mg, 0.122 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **55** as a white solid (56.0 mg, 0.109 mmol, 91%).

Rf: 0.66 (5% MeOH in DCM); ¹H NMR (600 MHz, CDCl₃) δ –0.01 (s, 9H, CH₂Si(CH₃)₃), 0.97 (t, *J* = 8.1 Hz, 2H, CH₂CH₂Si), 1.31 (t, *J* = 7.1 Hz, 3H, OCH₂CH₃), 3.60 (t, *J* = 8.0 Hz, 2H, OCH₂CH₂Si), 3.66 (q, *J* = 7.0 Hz, 2H, OCH₂CH₃), 3.84 (t, *J* = 4.9 Hz, 2H, OCH₂CH₂O), 4.17 (s, 3H, ArOCH₃), 4.29 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 5.68 (s, 2H, NCH₂O), 5.80 (d, *J* = 10.5 Hz, 1H, CCH=CH₂), 6.36 (dd, *J* = 16.9, 10.1 Hz, 1H, CCH=CH₂), 6.46 (d, *J* = 16.8 Hz, 1H, CCH=CH₂), 7.04 (d, *J* = 8.4 Hz, 1H, ArH), 7.36 (s, 1H, ArH), 7.43 (dd, *J* = 8.4, 2.1 Hz, 1H, ArH), 8.30 (s, 1H, ArH), 8.54 (s, 1H, ArH), 8.90 (s, 1H, ArNH) ppm; ¹³C NMR (151 MHz, CDCl₃) δ –1.3, 15.3, 17.8, 53.9, 66.3, 66.9, 68.8, 69.9, 73.2, 103.5, 113.7, 117.7, 120.6, 123.3, 123.7, 127.1, 128.2, 131.9, 133.2, 151.2, 153.0, 154.6, 163.6, 166.6 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₆H₃₇N₄O₅Si: 513.2533; found: 513.2537.

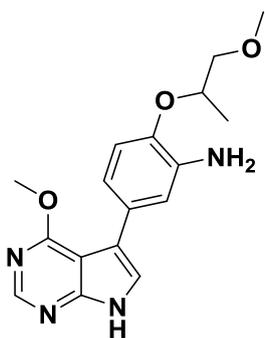
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

5-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-(2-methoxyethoxy)aniline (56)

Compound **56** was prepared according to the general procedure for *N*-SEM deprotection.

5-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-(2-methoxyethoxy)aniline (66.0 mg, 0.148 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **56** as a white solid (37.0 mg, 0.117 mmol, 78%).

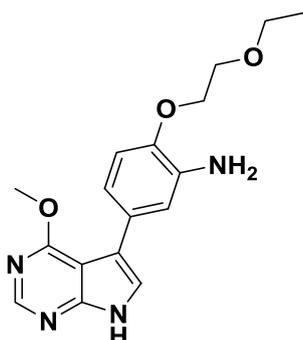
Rf: 0.63 (10% MeOH in DCM); ¹H NMR (600 MHz, CDCl₃) δ 3.39 (s, 3H, CH₂OCH₃), 3.71 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 4.01 (s, 3H, ArOCH₃), 4.12 (t, *J* = 4.8 Hz, 2H, OCH₂CH₂O), 6.78 (d, *J* = 8.2 Hz, 1H, ArH), 6.93 (dd, *J* = 8.2, 2.0 Hz, 1H, ArH), 6.98 – 6.95 (m, 1H, ArH), 7.11 (s, 1H, ArH), 8.41 (s, 1H, ArH), 11.35 (s, 1H, ArNH) ppm; ¹³C NMR (151 MHz, CDCl₃) δ 53.8, 59.2, 68.4, 71.3, 103.4, 112.7, 116.2, 117.6, 119.2, 120.5, 128.0, 136.6, 145.7, 150.6, 153.1, 163.7 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₆H₁₉N₄O₃: 315.1457; found: 315.1458.

5-(4-Methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]aniline (57)

Compound **57** was prepared according to the general procedure for *N*-SEM deprotection.

5-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]aniline (60.0 mg, 0.131 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **57** as a white solid (33.0 mg, 0.100 mmol, 77%).

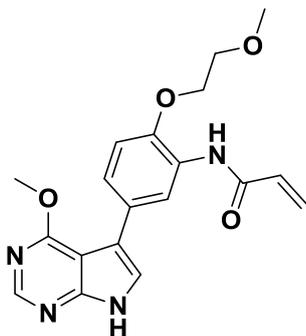
Rf: 0.65 (10% MeOH in DCM); ¹H NMR (600 MHz, CDCl₃) δ 1.36 (d, *J* = 6.3 Hz, 3H, CH₃CHCH₂), 3.43 (s, 3H, CHCH₂OCH₃), 3.53 – 3.50 (m, 1H, CHCH₂OCH₃), 3.64 – 3.60 (m, 1H, CHCH₂OCH₃), 4.08 (s, 3H, ArOCH₃), 4.11 (br s, 2H, ArNH₂), 4.48 – 4.44 (m, 1H, CH₃CHCH₂), 6.92 (d, *J* = 8.4 Hz, 1H, ArH), 7.03 (dd, *J* = 8.4, 1.9 Hz, 1H, ArH), 7.11 (d, *J* = 1.9 Hz, 1H, ArH), 7.19 (s, 1H, ArH), 8.47 (s, 1H, ArH), 11.42 (s, 1H, ArNH) ppm; ¹³C NMR (151 MHz, CDCl₃) δ 17.4, 53.8, 59.4, 75.1, 76.2, 103.4, 115.9, 117.0, 117.5, 119.8, 120.7, 128.5, 136.7, 144.9, 150.6, 153.0, 163.7 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₇H₂₁N₄O₃: 329.1614; found: 329.1620.

2-(2-Ethoxyethoxy)-5-(4-methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (58)

Compound **58** was prepared according to the general procedure for *N*-SEM deprotection. 2-(2-Ethoxyethoxy)-5-(4-methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl)aniline (60.0 mg, 0.131 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **58** as a white solid (32.0 mg, 97.5 μmol, 75%).

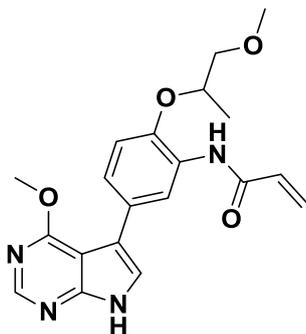
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

Rf: 0.63 (10% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 1.26 (t, $J = 7.0$ Hz, 3H, OCH_2OCH_3), 3.62 (q, $J = 7.0$ Hz, 2H, OCH_2OCH_3), 3.83 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.10 (s, 3H, ArOCH_3), 4.21 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 6.87 (d, $J = 8.2$ Hz, 1H, ArH), 7.02 (dd, $J = 8.2, 2.0$ Hz, 1H, ArH), 7.07 – 7.05 (m, 1H, ArH), 7.20 (s, 1H, ArH), 8.49 (s, 1H ArH), 11.41 (s, 1H, ArNH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 15.3, 53.8, 66.9, 68.7, 69.2, 103.4, 112.8, 116.4, 117.6, 119.4, 120.5, 128.0, 136.3, 145.8, 150.6, 153.1, 163.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{17}\text{H}_{21}\text{N}_4\text{O}_3$: 329.1614; found: 329.1613.

***N*-[5-(4-Methoxy-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)-2-(2-methoxyethoxy)phenyl]acrylamide (59)**

Compound **59** was prepared according to the general procedure for *N*-SEM deprotection. *N*-[5-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)-2-(2-methoxyethoxy)phenyl]acrylamide (40.0 mg, 80.2 μmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **59** as a white solid (24.0 mg, 65.1 μmol , 81%).

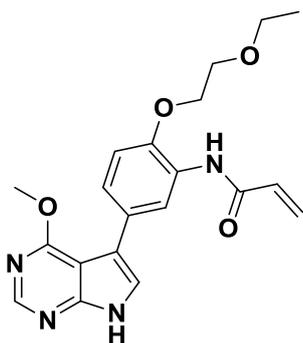
Rf: 0.60 (10% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 3.47 (s, 3H, CH_2OCH_3), 3.75 (t, $J = 4.5$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.15 (s, 3H, ArOCH_3), 4.23 (t, $J = 4.5$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.76 (d, $J = 10.4$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.32 (dd, $J = 16.8, 10.4$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.44 (d, $J = 16.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.00 (d, $J = 8.4$ Hz, 1H, ArH), 7.34 (s, 1H, ArH), 7.37 (dd, $J = 8.4, 2.0$ Hz, 1H, ArH), 8.34 (s, 1H, ArH), 8.50 (s, 1H, ArH), 8.85 (s, 1H, ArH), 11.26 (s, 1H, ArNH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 54.1, 59.2, 70.1, 71.0, 103.6, 114.3, 117.5, 120.7, 121.3, 124.5, 127.3, 128.6, 129.2, 131.9, 146.3, 150.0, 151.9, 163.5, 164.0 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{21}\text{N}_4\text{O}_4$: 369.1563; found: 369.1560.

***N*-[5-(4-Methoxy-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]phenyl]acrylamide (60)**

Compound **60** was prepared according to the general procedure for *N*-SEM deprotection. *N*-[5-(4-Methoxy-7-[[2-(trimethylsilyl)ethoxy]methyl]-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)-2-[(1-methoxypropan-2-yl)oxy]phenyl]acrylamide (42.0 mg, 81.2 μmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **60** as a white solid (24.0 mg, 63.3 μmol , 78%).

Rf: 0.61 (10% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 1.39 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 3.47 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.54 – 3.51 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.62 – 3.58 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 4.15 (s, 3H, ArOCH_3), 4.32 – 4.29 (m, 1H, CH_3CHCH_2), 5.75 (d, $J = 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.31 (dd, $J = 16.8, 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.44 (d, $J = 16.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.05 (d, $J = 8.4$ Hz, 1H, ArH), 7.36 (s, 1H, ArH), 7.37 (dd, $J = 8.4, 2.1$ Hz, 1H, ArH), 8.50 (s, 1H, ArH), 8.72 (s, 1H, ArNH), 8.90 (s, 1H, ArNH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 17.1, 54.1, 59.4, 76.0, 77.7, 103.5, 117.4, 118.3, 120.8, 121.4, 124.3, 127.2, 129.5, 130.7, 131.8, 145.6, 150.1, 152.2, 163.6, 163.9 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{20}\text{H}_{23}\text{N}_4\text{O}_4$: 383.1719; found: 383.1721.

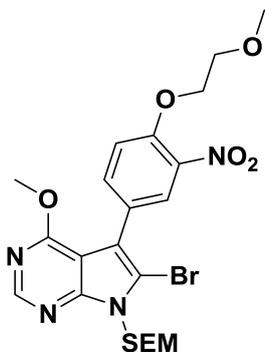
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-[2-(2-Ethoxyethoxy)-5-(4-methoxy-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]acrylamide (61)

Compound **61** was prepared according to the general procedure for *N*-SEM deprotection.

N-[2-(2-Ethoxyethoxy)-5-(4-methoxy-7-[2-(trimethylsilyl)ethoxy]methyl])-7H-pyrrolo[2,3-d]pyrimidin-5-yl]phenyl]acrylamide (46.0 mg, 89.7 μ mol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **61** as a white solid (27.0 mg, 70.8 μ mol, 79%).

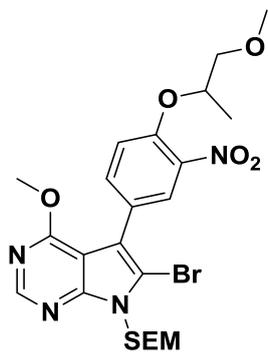
Rf: 0.59 (10% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 1.27 (t, J = 7.0 Hz, 3H, OCH_2CH_3), 3.62 (q, J = 7.0 Hz, 2H, OCH_2CH_3), 3.80 (t, J = 4.6 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.16 (s, 3H, ArOCH_3), 4.25 (t, J = 4.6 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.76 (d, J = 10.2 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.33 (dd, J = 16.8, 10.2 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.43 (d, J = 16.8 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.99 (d, J = 8.4 Hz, 1H, *ArH*), 7.33 (s, 1H, *ArH*), 7.36 (dd, J = 8.4, 1.9 Hz, 1H, *ArH*), 8.26 (s, 1H, *ArH*), 8.50 (s, 1H, *ArH*), 8.84 (s, 1H, *ArH*), 11.12 (s, 1H, *ArNH*) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 15.3, 54.2, 66.9, 68.8, 69.9, 103.7, 113.7, 117.7, 120.7, 121.3, 124.5, 127.3, 128.2, 129.0, 131.9, 146.4, 149.7, 151.4, 163.5, 164.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{20}\text{H}_{23}\text{N}_4\text{O}_4$: 383.1719; found: 383.1720.

6-Bromo-4-methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-7-[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (62)

Compound **62** was prepared according to the general procedure for pyrrolopyrimidine derivative bromination. 4-Methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-7-[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (130 mg, 0.274 mmol) was brominated and purification was achieved using flash column chromatography with elution gradient of 0 – 30% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **62** as a yellow solid (114 mg, 0.206 mmol, 75%).

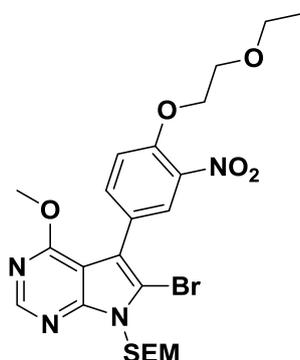
Rf: 0.41 (30% EtOAc in PE); ^1H NMR (400 MHz, CDCl_3) δ -0.04 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.95 (t, J = 8.2 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.49 (s, 3H, CH_2OCH_3), 3.64 (t, J = 8.2 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.84 (t, J = 4.7 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.03 (s, 3H, ArOCH_3), 4.32 (t, J = 4.7 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.76 (s, 2H, NCH_2O), 7.17 (d, J = 8.7 Hz, 1H, *ArH*), 7.71 (dd, J = 8.7, 2.3 Hz, 1H, *ArH*), 8.06 (d, J = 2.2 Hz, 1H, *ArH*), 8.50 (s, 1H, *ArH*) ppm; ^{13}C NMR (101 MHz, CDCl_3) δ -1.3, 17.9, 54.3, 59.7, 67.0, 69.8, 70.8, 72.2, 104.4, 111.9, 114.4, 114.4, 125.4, 128.1, 136.0, 139.6, 151.6, 151.6, 152.7, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{22}\text{H}_{30}\text{N}_4\text{O}_6\text{SiBr}$: 553.1118; found: 553.1113.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

6-Bromo-4-methoxy-5-{4-[(1-methoxypropan-2-yl)oxy]-3-nitrophenyl}-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (63)

Compound **63** was prepared according to the general procedure for pyrrolopyrimidine derivative bromination. 4-Methoxy-5-{4-[(1-methoxypropan-2-yl)oxy]-3-nitrophenyl}-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (288 mg, 0.589 mmol) was brominated and purification was achieved using flash column chromatography with elution gradient of 0 – 30% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **63** as a yellow solid (237 mg, 0.418 mmol, 71%).

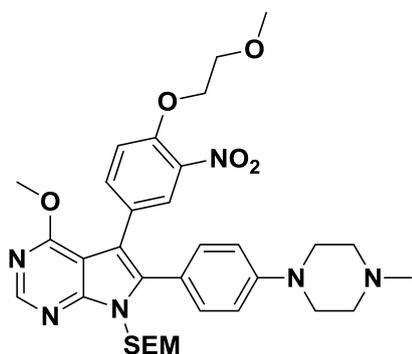
Rf: 0.45 (30% EtOAc in PE); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 0.00 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 1.00 (t, $J = 8.3$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.46 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 3.48 (s, 2H, $\text{CHCH}_2\text{OCH}_3$), 3.63 – 3.59 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.68 (t, 2H, $J = 8.3$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.75 – 3.70 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 4.07 (s, 3H, ArOCH_3), 4.79 – 4.74 (m, 1H, CH_3CHCH_2), 5.79 (s, 2H, NCH_2O), 7.27 (d, $J = 8.8$ Hz, 1H, ArH), 7.74 (dd, $J = 8.8$, 2.2 Hz, 1H, ArH), 8.06 (d, $J = 2.2$ Hz, 1H, ArH), 8.53 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ –1.3, 17.1, 17.9, 54.1, 59.6, 66.9, 72.1, 76.0, 76.3, 104.3, 111.8, 114.3, 115.8, 125.2, 127.9, 135.7, 140.5, 150.9, 151.7, 152.8, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{32}\text{N}_4\text{O}_6\text{SiBr}$: 567.1275; found: 567.1280.

6-Bromo-5-[4-(2-ethoxyethoxy)-3-nitrophenyl]-4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (64)

Compound **64** was prepared according to the general procedure for pyrrolopyrimidine derivative bromination. 5-[4-(2-Ethoxyethoxy)-3-nitrophenyl]-4-methoxy-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (307 mg, 0.628 mmol) was brominated and purification was achieved using flash column chromatography with elution gradient of 0 – 30% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **64** as a yellow solid (257 mg, 0.453 mmol, 72%).

Rf: 0.46 (30% EtOAc in PE); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ –0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.94 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.22 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 3.60 (t, $J = 8.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.63 (q, $J = 7.0$ Hz, 2H), 3.86 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.00 (s, 3H, ArOCH_3), 4.31 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.73 (s, 2H, NCH_2O), 7.17 (d, $J = 8.7$ Hz, 1H, ArH), 7.70 (dd, $J = 8.7$, 2.2 Hz, 1H, ArH), 8.05 (d, $J = 2.2$ Hz, 1H, ArH), 8.47 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ –1.4, 15.3, 17.8, 54.1, 66.9, 67.3, 68.7, 69.8, 72.1, 104.3, 111.7, 114.2, 114.3, 125.2, 128.0, 136.0, 139.5, 151.6, 151.7, 152.8, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{23}\text{H}_{32}\text{N}_4\text{O}_6\text{SiBr}$: 567.1275; found: 567.1272.

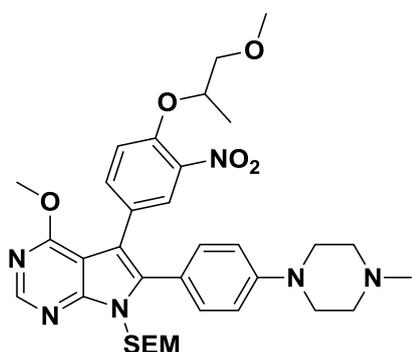
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

4-Methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (65)

Compound **65** was prepared according to the general procedure for Suzuki-Miyaura cross coupling reaction. 6-Bromo-4-methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (314 mg, 0.567 mmol) and 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine (205 mg, 0.678 mmol) were coupled and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **65** as a yellow

foam (335 mg, 0.516 mmol, 91%).

Rf: 0.48 (5% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ -0.04 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.94 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 2.49 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.76 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.40 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.46 (s, 3H, CH_2OCH_3), 3.66 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.79 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.04 (s, 3H, ArOCH_3), 4.23 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.51 (s, 2H, NCH_2O), 6.89 – 6.86 (m, 2H, $2 \times \text{ArH}$), 6.93 (dd [app. t], $J = 8.6$ Hz, 1H, ArH), 7.28 (d, $J = 8.6$ Hz, 2H, $2 \times \text{ArH}$), 7.32 (dd, $J = 8.6, 2.1$ Hz, 1H, ArH), 7.88 (d, $J = 2.1$ Hz, 1H, ArH), 8.51 (s, 1H, ArH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ -1.3, 18.1, 47.6, 53.9, 54.5, 54.7, 59.7, 69.5, 70.8, 71.0, 104.1, 111.5, 114.0, 115.6, 116.7, 127.2, 128.4, 129.4, 132.5, 136.5, 137.1, 139.4, 150.6, 151.3, 153.1, 163.0 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{33}\text{H}_{45}\text{N}_6\text{O}_6\text{Si}$: 649.3170; found: 649.3170.

4-Methoxy-5-[4-[(1-methoxypropan-2-yl)oxy]-3-nitrophenyl]-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (66)

Compound **66** was prepared according to the general procedure for Suzuki-Miyaura cross coupling reaction. 6-Bromo-4-methoxy-5-[4-[(1-methoxypropan-2-yl)oxy]-3-nitrophenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidine (240 mg, 0.423 mmol) and 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine (154 mg, 0.510 mmol) were coupled and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford

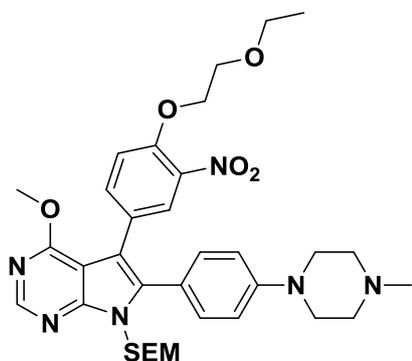
compound (HL788 **66**) as a yellow foam (264 mg, 0.398 mmol, 94%).

Rf: 0.47 (5% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ -0.04 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.94 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.36 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 2.50 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.54 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.37 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.41 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.53 – 3.50 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.67 – 3.61 (m, 3H, $\text{CHCH}_2\text{OCH}_3$ and $\text{OCH}_2\text{CH}_2\text{Si}$), 4.05 (s, 3H, ArOCH_3), 4.64 – 4.60 (m, 1H, CH_3CHCH_2), 5.51 (s, 2H, NCH_2O), 6.90 – 6.87 (m, 3H, $3 \times \text{ArH}$), 6.97 (d, $J = 8.8$ Hz, 1H, ArH), 7.31 – 7.29 (m, 2H, $2 \times \text{ArH}$), 7.83 (d, $J = 1.9$ Hz, 1H, ArH), 8.52 (s, 1H, ArH) ppm;

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

^{13}C NMR (126 MHz, CDCl_3) δ -1.3, 17.0, 18.1, 47.6, 48.4, 53.9, 54.7, 59.6, 66.8, 70.9, 76.0, 76.1, 104.1, 111.5, 115.6, 116.7, 120.7, 127.0, 128.2, 129.4, 132.5, 136.1, 137.0, 140.4, 149.8, 151.2, 153.1, 163.0 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{34}\text{H}_{47}\text{N}_6\text{O}_6\text{Si}$: 663.3326; found: 663.3327.

5-[4-(2-Ethoxyethoxy)-3-nitrophenyl]-4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidine (67)

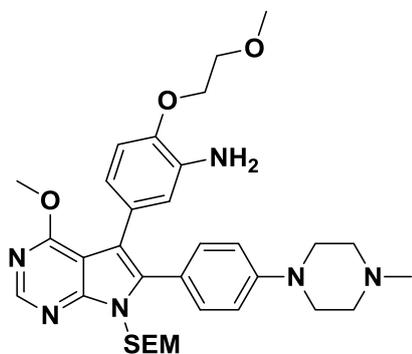


Compound **67** was prepared according to the general procedure for Suzuki-Miyaura cross coupling reaction. 6-Bromo-5-[4-(2-ethoxyethoxy)-3-nitrophenyl]-4-methoxy-7-[2-(trimethylsilyl)ethoxy]methyl-7H-pyrrolo[2,3-d]pyrimidine (240 mg, 0.423 mmol) and 1-methyl-4-[4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]piperazine (154 mg, 0.510 mmol) were coupled and purification was achieved using flash column chromatography with elution gradient of 0 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **67** as a yellow

foam (255 mg, 0.385 mmol, 91%).

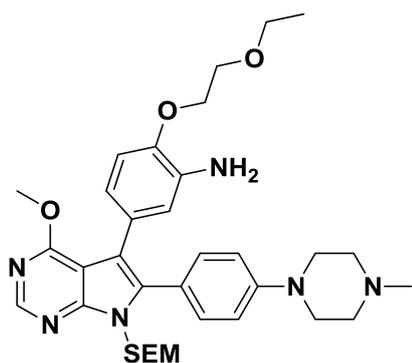
Rf: 0.48 (5% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ -0.04 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.95 (t, J = 8.4 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.23 (t, J = 7.0 Hz, 3H, OCH_2CH_3), 2.61 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.91 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.46 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.62 (q, J = 7.0 Hz, 2H, OCH_2CH_3), 3.66 (t, J = 8.4 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.83 (t, J = 4.8 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.04 (s, 3H, ArOCH_3), 4.23 (t, J = 4.8 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.51 (s, 2H, NCH_2O), 6.87 (d, J = 8.7 Hz, 2H, $2 \times \text{ArH}$), 6.94 (d, J = 8.1 Hz, 1H, ArH), 7.30 – 7.28 (m, 2H, $2 \times \text{ArH}$), 7.34 (dd, J = 8.7, 2.0 Hz, 1H, ArH), 7.83 (d, J = 2.0 Hz, 1H, ArH), 8.52 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ -1.3, 15.3, 18.1, 47.1, 48.0, 53.9, 54.3, 66.8, 67.3, 68.7, 69.6, 71.0, 104.1, 111.6, 114.1, 115.8, 117.0, 127.0, 128.2, 129.5, 132.6, 136.5, 136.9, 139.4, 150.7, 151.3, 153.1, 163.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{34}\text{H}_{47}\text{N}_6\text{O}_6\text{Si}$: 663.3326; found: 663.3325.

5-[4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl]-2-(2-methoxyethoxy)aniline (68)



Compound **68** was prepared according to the general procedure for aromatic nitro derivative reduction. 4-Methoxy-5-[4-(2-methoxyethoxy)-3-nitrophenyl]-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[2-(trimethylsilyl)ethoxy]methyl-7H-pyrrolo[2,3-d]pyrimidine (350 mg, 0.539 mmol) was reduced and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **68** as a white foam (280 mg, 0.453 mmol, 84%).

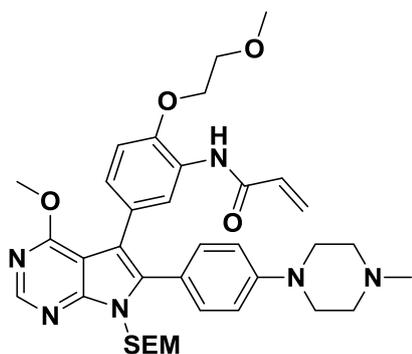
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors



Rf: 0.43 (5% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ -0.02 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.96 (t, J = 8.4 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.26 (t, J = 7.0 Hz, 3H, OCH_2CH_3), 2.39 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.62 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.29 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.62 (q, J = 7.0 Hz, 2H, OCH_2CH_3), 3.67 (t, J = 8.4 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.78 (br s, 2H, ArNH_2), 3.82 (t, J = 4.8 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.03 (s, 3H, ArOCH_3), 4.15 (t, J = 4.8 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.53 (s, 2H, NCH_2O), 6.62 (dd, J = 8.2, 1.9 Hz, 1H, ArH), 6.70 (d, J = 8.3 Hz, 1H, ArH), 6.72 (d, J = 1.9 Hz, 1H, ArH), 6.88 – 6.85 (m, 2H, $2 \times \text{ArH}$),

7.32 – 7.30 (m, 2H, $2 \times \text{ArH}$), 8.50 (s, 1H, ArH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ -1.3, 15.3, 18.0, 46.1, 48.2, 53.7, 55.0, 66.5, 66.8, 68.1, 69.2, 70.9, 104.5, 111.6, 114.0, 115.0, 118.4, 121.1, 121.5, 127.4, 132.3, 135.7, 136.2, 145.2, 150.6, 150.6, 152.9, 163.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{34}\text{H}_{49}\text{N}_6\text{O}_4\text{Si}$: 633.3585; found: 633.3586.

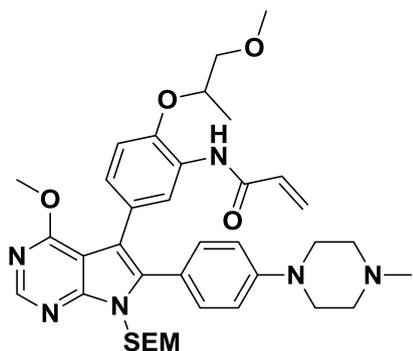
***N*-{5-[4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl]-2-(2-methoxyethoxy)phenyl}acrylamide (71)**



Compound **71** was prepared according to the general procedure for aniline derivative acylation. 5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl}-2-(2-methoxyethoxy)aniline (200 mg, 0.323 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **71** as a white solid (172 mg, 0.256 mmol, 79%).

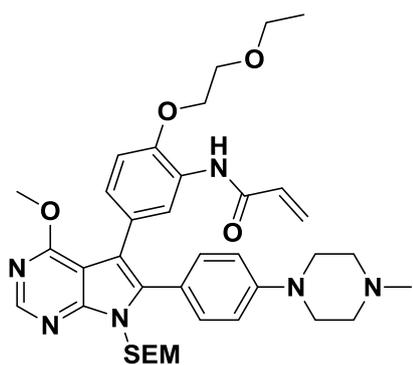
Rf: 0.59 (5% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 0.00 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.98 (t, J = 8.4 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 2.38 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.61 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.28 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.47 (s, 3H, CH_2OCH_3), 3.68 (t, J = 8.4 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.74 (t, J = 4.7 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.05 (s, 3H, ArOCH_3), 4.17 (t, J = 4.7 Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.55 (s, 2H, NCH_2O), 5.73 (d, J = 10.2 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.30 (dd, J = 16.8, 10.2 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.39 (d, J = 16.8 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.77 (d, J = 8.3 Hz, 1H, ArH), 6.86 – 6.82 (m, 1H, ArH), 6.87 (s, 1H, ArH), 6.89 (s, 1H, ArH), 7.73 (s, 1H, ArH), 7.75 (s, 1H, ArH), 8.24 (s, 1H, ArH), 8.52 (s, 1H, ArH), 8.65 (s, 1H, ArNH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ -1.4, 18.0, 46.0, 48.2, 48.9, 53.6, 55.0, 59.1, 66.4, 69.4, 70.9, 104.5, 112.7, 113.8, 115.1, 116.1, 119.8, 120.9, 123.4, 126.4, 126.7, 128.3, 129.1, 132.0, 132.3, 136.4, 150.6, 150.7, 152.8, 163.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{36}\text{H}_{49}\text{N}_6\text{O}_5\text{Si}$: 673.3534; found: 673.3534.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-(5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(1-methoxypropan-2-yl)phenyl)acrylamide (72)

Compound **72** was prepared according to the general procedure for aniline derivative acylation. 5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(1-methoxypropan-2-yl)aniline (100 mg, 0.158 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **72** as a white solid (81.0 mg, 0.118 mmol, 75%).

Rf: 0.60 (5% MeOH in DCM); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ –0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.93 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.32 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 2.36 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.60 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.26 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.43 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.49 – 3.46 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.57 – 3.53 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.62 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 4.30 – 4.23 (m, 1H, CH_3CHCH_2), 5.49 (s, 2H, NCH_2O), 5.68 (d, $J = 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.25 (dd, $J = 16.9, 10.2$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.35 (dd, $J = 16.9, 0.9$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.80 – 6.76 (m, 2H, $2 \times \text{ArH}$), 6.82 (s, 1H, ArH), 6.84 (s, 1H, ArH), 7.29 (s, 1H, ArH), 7.30 (s, 1H, ArH), 8.47 (s, 1H, ArH), 8.54 (s, 1H, ArH), 8.63 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ –1.3, 17.1, 18.1, 46.0, 48.2, 53.7, 55.0, 59.3, 66.5, 70.9, 76.0, 77.0, 104.6, 113.9, 115.2, 116.8, 121.2, 123.6, 126.4, 126.7, 129.3, 129.8, 132.0, 132.5, 136.5, 145.0, 150.6, 150.8, 152.9, 163.2, 163.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{37}\text{H}_{51}\text{N}_6\text{O}_5\text{Si}$: 687.3690; found: 687.3691.

N-[2-(2-Ethoxyethoxy)-5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl]acrylamide (73)

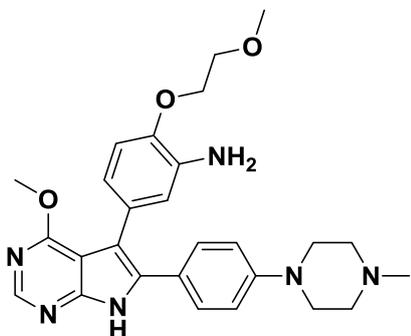
Compound **73** was prepared according to the general procedure for aniline derivative acylation. 2-(2-Ethoxyethoxy)-5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-{[2-(trimethylsilyl)ethoxy]methyl}-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (120 mg, 0.190 mmol) was acylated and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **73** as a white solid (94.0 mg, 0.137 mmol, 72%).

Rf: 0.61 (5% MeOH in DCM); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ –0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.92 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 1.23 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.36 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.60 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.25 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.58 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.62 (t, $J = 8.4$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{Si}$), 3.74 (t, $J = 4.6$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.99 (s, 3H, ArOCH_3), 4.14 (t, $J = 4.6$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.49 (s, 2H, NCH_2O), 5.68 (d, $J = 10.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.25 (dd, $J = 16.8, 10.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.33 (dd, $J = 16.8, 1.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.72 (d, $J = 8.3$ Hz, 1H, ArH), 6.79 (d, $J = 7.7$ Hz, 1H, ArH), 6.82 (s, 1H, ArH), 6.83 (s, 1H, ArH), 7.28 (s, 1H, ArH), 7.29 (s, 1H, ArH), 8.11 (s, 1H, ArH), 8.47 (s, 1H, ArH), 8.57 (s, 1H, ArNH) ppm;

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

^{13}C NMR (151 MHz, CDCl_3) δ -1.3, 15.3, 18.1, 46.0, 48.2, 53.6, 55.0, 66.5, 66.8, 68.8, 69.3, 70.9, 104.6, 112.3, 113.9, 115.2, 121.1, 123.4, 126.5, 126.7, 128.0, 128.1, 132.1, 132.4, 136.4, 145.8, 150.6, 150.8, 152.8, 163.0, 163.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{37}\text{H}_{51}\text{N}_6\text{O}_5\text{Si}$: 687.3690; found: 687.3693.

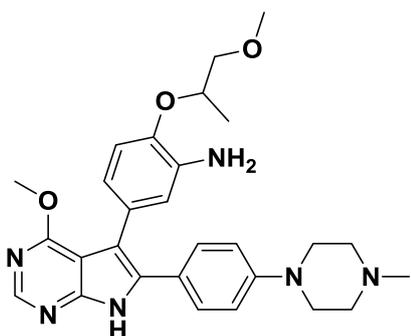
5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(2-methoxyethoxy)aniline (74)



Compound **74** was prepared according to the general procedure for *N*-SEM deprotection. 5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[(2-(trimethylsilyl)ethoxy)methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(2-methoxyethoxy)aniline (101 mg, 0.163 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **74** as a white solid (60.0 mg, 0.123 mmol, 73%).

Rf: 0.31 (10% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 2.38 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.59 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.21 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.47 (s, 3H, CH_2OCH_3), 3.80 (t, $J = 4.6$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.84 (br s, 2H, ArNH_2), 3.90 (s, 3H, ArOCH_3), 4.19 (t, $J = 4.6$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 6.80 – 6.75 (m, 5H, $5 \times \text{ArH}$), 7.28 – 7.24 (m, 2H, $2 \times \text{ArH}$), 8.29 (s, 1H, ArH), 13.19 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 46.0, 48.0, 53.6, 55.0, 59.3, 67.9, 71.4, 106.2, 111.3, 111.7, 115.3, 118.5, 121.5, 122.8, 128.0, 129.2, 133.6, 136.0, 145.4, 149.7, 150.2, 152.2, 163.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{33}\text{N}_6\text{O}_3$: 489.2614; found: 489.2610.

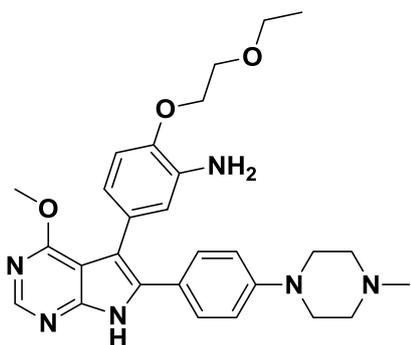
5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(1-methoxypropan-2-yl)aniline (75)



Compound **75** was prepared according to the general procedure for *N*-SEM deprotection. 5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[(2-(trimethylsilyl)ethoxy)methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(1-methoxypropan-2-yl)aniline (70.0 mg, 0.111 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **75** as a white solid (41.0 mg, 81.5 μmol , 74%).

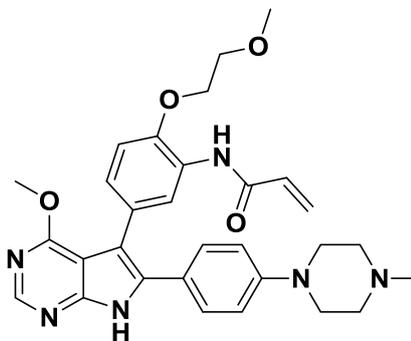
Rf: 0.32 (10% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 1.37 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 2.39 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.61 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.25 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.44 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.53 – 3.49 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.66 – 3.61 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.89 (br s, 2H, ArNH_2), 3.94 (s, 3H, ArOCH_3), 4.52 – 4.45 (m, 1H, CH_3CHCH_2), 6.73 (d, $J = 8.1$ Hz, 1H, ArH), 6.79 (s, 1H, ArH), 6.85 – 6.80 (m, $J = 8.2, 5.3$ Hz, 3H, $3 \times \text{ArH}$), 7.33 (s, 1H, ArH), 7.35 (s, 1H, ArH), 8.32 (s, 1H, ArH), 13.21 (s, 1H, ArNH) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 17.4, 46.0, 48.1, 53.7, 55.0, 59.4, 74.5, 76.3, 106.2, 111.4, 115.0, 115.3, 118.7, 121.4, 122.9, 128.5, 129.3, 133.6, 137.3, 144.2, 149.7, 150.2, 152.2, 163.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{35}\text{N}_6\text{O}_3$: 503.2771; found: 503.2770.

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

2-(2-Ethoxyethoxy)-5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (76)

Compound **76** was prepared according to the general procedure for *N*-SEM deprotection. 2-(2-Ethoxyethoxy)-5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}aniline (70.0 mg, 0.111 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **76** as a white solid (40.0 mg, 79.6 μ mol, 73%).

Rf: 0.32 (10% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 1.26 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.41 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.63 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.25 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.63 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.84 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.89 (br s, 2H, ArNH_2), 3.93 (s, 3H, ArOCH_3), 4.19 (t, $J = 4.8$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 6.76 (d, $J = 8.2$ Hz, 1H, ArH), 6.83 – 6.77 (m, 4H, $4 \times \text{ArH}$), 7.34 – 7.28 (m, 2H, $2 \times \text{ArH}$), 8.32 (s, 1H, ArH), 13.08 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 15.4, 46.0, 48.0, 53.7, 55.0, 66.9, 68.2, 69.3, 106.2, 111.4, 111.9, 115.4, 118.4, 121.5, 123.0, 128.0, 129.2, 133.6, 136.1, 145.5, 149.7, 150.2, 152.2, 163.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{35}\text{N}_6\text{O}_3$: 503.2771; found: 503.2773.

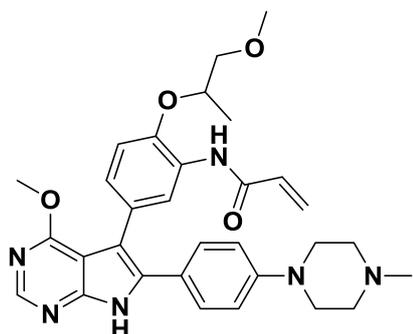
N-[5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(2-methoxyethoxy)phenyl]acrylamide (77)

Compound **77** was prepared according to the general procedure for *N*-SEM deprotection. *N*-[5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[[2-(trimethylsilyl)ethoxy]methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(2-methoxyethoxy)phenyl]acrylamide (171 mg, 0.254 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **77** as a white solid (99.0 mg, 0.182 mmol,

72%).

Rf: 0.37 (10% MeOH in DCM); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 2.38 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.62 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.28 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.47 (s, 3H, CH_2OCH_3), 3.76 (t, $J = 4.3$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 4.00 (s, 3H, ArOCH_3), 4.21 (t, $J = 4.3$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.72 (d, $J = 10.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.30 (dd, $J = 16.8$, 10.0 Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.38 (d, $J = 16.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.88 – 6.83 (m, 3H, $3 \times \text{ArH}$), 6.99 (d, $J = 8.0$ Hz, 1H, ArH), 7.39 (s, 1H, ArH), 7.41 (s, 1H, ArH), 8.26 (s, 1H, ArH), 8.37 (s, 1H, ArH), 8.68 (s, 1H, ArNH), 13.11 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 46.0, 48.2, 53.7, 55.0, 59.2, 69.5, 71.0, 106.2, 111.2, 113.1, 115.5, 123.0, 123.5, 126.7, 127.0, 128.5, 129.0, 129.5, 132.0, 133.9, 146.1, 149.9, 150.4, 152.2, 163.2, 163.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{35}\text{N}_6\text{O}_4$: 543.2720; found: 543.2717.

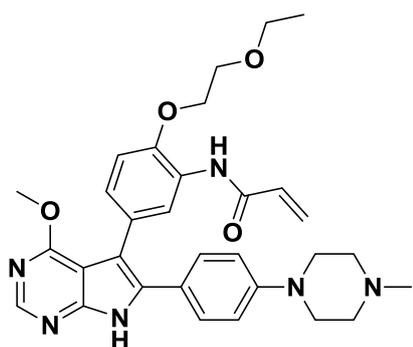
Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

N-[5-{4-Methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(1-methoxypropan-2-yl)phenyl]acrylamide (78)

(52.0 mg, 93.4 μ mol, 79%).

Compound **78** was prepared according to the general procedure for *N*-SEM deprotection. *N*-[5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[(2-(trimethylsilyl)ethoxy)methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}-2-(1-methoxypropan-2-yl)phenyl]acrylamide (81.0 mg, 0.118 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **78** as a white solid

Rf: 0.39 (10% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 1.39 (d, $J = 6.3$ Hz, 3H, CH_3CHCH_2), 2.39 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.63 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.29 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.48 (s, 3H, $\text{CHCH}_2\text{OCH}_3$), 3.56 – 3.52 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 3.64 – 3.59 (m, 1H, $\text{CHCH}_2\text{OCH}_3$), 4.02 (s, 3H, ArOCH_3), 4.38 – 4.32 (m, 1H, CH_3CHCH_2), 5.73 (d, $J = 10.1$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.31 (dd, $J = 16.9, 10.1$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.41 (d, $J = 16.9$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.88 (s, 1H, *ArH*), 6.89 (s, 1H, *ArH*), 6.92 (d, $J = 8.3$ Hz, 1H, *ArH*), 6.99 (d, $J = 8.0$ Hz, 1H, *ArH*), 7.43 (s, 1H, *ArH*), 7.45 (s, 1H, *ArH*), 8.40 (s, 1H, *ArH*), 8.66 (s, 1H, *ArH*), 8.75 (s, 1H, *ArNH*), 13.28 (s, 1H, *ArNH*) ppm; ^{13}C NMR (126 MHz, CDCl_3) δ 17.1, 46.0, 48.2, 53.7, 55.0, 59.3, 76.0, 106.2, 111.2, 115.4, 117.3, 122.9, 123.6, 126.6, 126.9, 129.5, 129.9, 130.2, 131.9, 134.0, 145.3, 149.8, 150.4, 152.2, 163.3, 163.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{31}\text{H}_{37}\text{N}_6\text{O}_4$: 557.2876; found: 557.2881.

N-[2-(2-Ethoxyethoxy)-5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl]acrylamide (79)

Compound **79** was prepared according to the general procedure for *N*-SEM deprotection.

N-[2-(2-Ethoxyethoxy)-5-{4-methoxy-6-[4-(4-methylpiperazin-1-yl)phenyl]-7-[(2-(trimethylsilyl)ethoxy)methyl]-7H-pyrrolo[2,3-d]pyrimidin-5-yl}phenyl]acrylamide (79.0 mg, 0.115 mmol) was deprotected and purification was achieved using flash column chromatography with elution gradient of 0 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **79** as a white solid (49.0 mg, 88.0 μ mol, 77%).

Rf: 0.38 (10% MeOH in DCM); ^1H NMR (500 MHz, CDCl_3) δ 1.26 (t, $J = 7.0$ Hz, 3H, OCH_2CH_3), 2.37 (s, 3H, $(\text{CH}_2)_2\text{NCH}_3$), 2.61 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.27 (br s, 4H, $2 \times \text{NCH}_2\text{CH}_2\text{N}$), 3.61 (q, $J = 7.0$ Hz, 2H, OCH_2CH_3), 3.79 (t, $J = 4.5$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 3.99 (s, 3H, ArOCH_3), 4.20 (t, $J = 4.5$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{O}$), 5.71 (d, $J = 10.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.30 (dd, $J = 16.8, 10.0$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.37 (d, $J = 16.8$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.88 – 6.81 (m, 3H, $3 \times \text{ArH}$), 6.98 (d, $J = 7.7$ Hz, 1H, *ArH*), 7.39 (s, 1H, *ArH*), 7.41 (s, 1H, *ArH*), 8.20 (s, 1H, *ArH*), 8.37 (s, 1H, *ArH*), 8.67 (s, 1H, *ArNH*), 13.22 (s, 1H, *ArNH*) ppm;

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

^{13}C NMR (126 MHz, CDCl_3) δ 15.3, 46.0, 48.1, 53.7, 54.9, 66.9, 68.9, 69.3, 106.2, 111.2, 112.6, 115.5, 122.9, 123.5, 126.7, 126.9, 128.3, 128.7, 129.5, 132.0, 133.9, 146.1, 149.8, 150.4, 152.2, 163.2, 163.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{31}\text{H}_{37}\text{N}_6\text{O}_4$: 557.2876; found: 557.2874.

Crystallisation of cSrc-T338M/S345C with compound 77

The inhibitor **77** was co-crystallized with cSrc-T338M/S345C by incubating 5 mg/mL protein with a threefold molar excess of inhibitor (10 mM DMSO stock) for 1 h at 4 °C to allow enzyme–inhibitor complex formation prior to crystallization. Crystals were grown using the hanging drop method at 20 °C after mixing a 1 μL protein-inhibitor solution with a 1 μL reservoir solution (10% ethylene glycol, 1 mM sodium chloride, pH 7.0). The data sets were collected at the PX10SA beamline of the Swiss Light Source (PSI, Villigen, Switzerland). All data sets were processed with XDS and scaled using XSCALE.³⁰

The complex crystal structures of cSrc-T338M/S345C in complex with compound **77** was solved by molecular replacement with PHASER using structure PDB: 3G5D as template.³¹ The cSrc molecule in the asymmetric unit was manually adjusted using the program COOT.³² The refinement was performed with Phenix 1.10.³³ Inhibitor topology files were generated using the Dundee PRODRG2 server and refined structures were validated with PROCHECK and the PDB validation server.³⁴ PyMOL was utilised for generating **Figure 4.11**. Data collection, structure refinement statistics, and further details for data collection are provided in **Table 4.10** below.

Table 4.10: Data collection and refinement statistics for cSrc-T338M/S345C in complex with **77**.

cSrc-T338M/S345C with 77			
Data Collection		Refinement	
Space Group	P1	Resolution [Å]	43.31 – 1.90
Cell Dimensions		No. reflections	57563
a, b, c [Å]	42.4, 63.3, 74.6	$R_{\text{work}} / R_{\text{free}}$	27 / 30
α, β, γ [°]	78.2, 90.7, 89.8	No. atoms	
Resolution [Å]	40.3-2.1 (2.2-2.1)	Protein	4127
R_{meas} [%]	7 (62.6)	Ligand/ion	76
$I / \sigma I$	6.5 (2.5)	Water	321
Completeness [%]	97.6 (97)	B-factors	
$\text{CC}_{1/2}$	99.8 (84.3)	Protein	50.72
Redundancy	2.3 (2.4)	Ligand/ion	54.38
		Water	46.61
		R.M.S. deviations	
		Bond lengths [Å]	0.006
		Bond angles [°]	0.817

4.11 References

1. P. Traxler, P. R. Allegrini, R. Brandt, J. Brueggen, R. Cozens, D. Fabbro, K. Grosios, H. A. Lane, P. McSheehy, J. Mestan, T. Meyer, C. Tang, M. Wartmann, J. Wood and G. Caravatti, *Cancer Research*, 2004, **64**, 4931-4941.
2. T. Ishikawa, M. Seto, H. Banno, Y. Kawakita, M. Oorui, T. Taniguchi, Y. Ohta, T. Tamura, A. Nakayama, H. Miki, H. Kamiguchi, T. Tanaka, N. Habuka, S. Sogabe, J. Yano, K. Aertgeerts and K. Kamiyama, *Journal of Medicinal Chemistry*, 2011, **54**, 8030-8050.
3. S. Sogabe, Y. Kawakita, S. Igaki, H. Iwata, H. Miki, D. R. Cary, T. Takagi, S. Takagi, Y. Ohta and T. Ishikawa, *ACS Medicinal Chemistry Letters*, 2013, **4**, 201-205.
4. P. A. Jackson, J. C. Widen, D. A. Harki and K. M. Brummond, *Journal of Medicinal Chemistry*, 2017, **60**, 839-885.
5. R. Lonsdale, J. Burgess, N. Colclough, N. L. Davies, E. M. Lenz, A. L. Orton and R. A. Ward, *Journal of Chemical Information and Modeling*, 2017, **57**, 3124-3137.
6. C. Bissantz, B. Kuhn and M. Stahl, *Journal of Medicinal Chemistry*, 2010, **53**, 5061-5084.
7. A. J. Parker, *Chemical Reviews*, 1969, **69**, 1-32.
8. K. A. Foster, C. G. Oster, M. M. Mayer, M. L. Avery and K. L. Audus, *Experimental Cell Research*, 1998, **243**, 359-366.
9. M. Warmuth, S. Kim, X.-j. Gu, G. Xia and F. Adrián, *Current Opinion in Oncology*, 2007, **19**, 55-60.
10. J. A. Blair, D. Rauh, C. Kung, C.-H. Yun, Q.-W. Fan, H. Rode, C. Zhang, M. J. Eck, W. A. Weiss and K. M. Shokat, *Nature Chemical Biology*, 2007, **3**, 229-238.
11. A. Michalczyk, S. Klüter, H. B. Rode, J. R. Simard, C. Grütter, M. Rabiller and D. Rauh, *Bioorganic & Medicinal Chemistry*, 2008, **16**, 3482-3488.
12. B. Lippa, G. S. Kauffman, J. Arcari, T. Kwan, J. Chen, W. Hungerford, S. Bhattacharya, X. Zhao, C. Williams, J. Xiao, L. Pustilnik, C. Su, J. D. Moyer, L. Ma, M. Campbell and S. Steyn, *Bioorganic & Medicinal Chemistry Letters*, 2007, **17**, 3081-3086.
13. Y. Kawakita, H. Banno, T. Ohashi, T. Tamura, T. Yusa, A. Nakayama, H. Miki, H. Iwata, H. Kamiguchi, T. Tanaka, N. Habuka, S. Sogabe, Y. Ohta and T. Ishikawa, *Journal of Medicinal Chemistry*, 2012, **55**, 3975-3991.
14. J. Kim, L. Wang, Y. Li, K. D. Becnel, K. M. Frey, S. J. Garforth, V. R. Prasad, R. F. Schinazi, D. C. Liotta and K. S. Anderson, *Bioorganic & Medicinal Chemistry Letters*, 2012, **22**, 4064-4067.
15. J. C. Reader, T. P. Matthews, S. Klair, K.-M. J. Cheung, J. Scanlon, N. Proisy, G. Addison, J. Ellard, N. Piton, S. Taylor, M. Cherry, M. Fisher, K. Boxall, S. Burns, M. I. Walton, I. M. Westwood, A. Hayes, P. Eve, M. Valenti, A. de Haven Brandon, G. Box, R. L. M. van Montfort, D. H. Williams, G. W. Aherne, F. I. Raynaud, S. A. Eccles, M. D. Garrett and I. Collins, *Journal of Medicinal Chemistry*, 2011, **54**, 8328-8342.
16. N. Sabat, S. Smoleń, P. Nauš, P. Perlíková, M. Cebová, L. Poštová Slavětínská and M. Hocek, *Synthesis*, 2017, **49**, 4623-4650.
17. L. Zhang, Y. Zhang, X. Li and L. Zhang, *Bioorganic & Medicinal Chemistry*, 2002, **10**, 907-912.
18. S. M. Maddox, C. J. Nalbandian, D. E. Smith and J. L. Gustafson, *Organic Letters*, 2015, **17**, 1042-1045.
19. A. B. Gamble, J. Garner, C. P. Gordon, S. M. J. O'Conner and P. A. Keller, *Synthetic Communications*, 2007, **37**, 2777-2786.
20. S. Jeon, M. Wang, L.-S. Tan, T. Cooper, M. Hamblin and L. Chiang, *Molecules*, 2013, **18**, 9603.
21. J. P. Davin, J.-C. Buffet, T. P. Spaniol and J. Okuda, *Dalton Transactions*, 2012, **41**, 12612-12618.
22. W. Shan, F. Meng, Y. Wu, F. Mao and X. Li, *Journal of Organometallic Chemistry*, 2011, **696**, 1687-1690.
23. Y. Wang, L. Deng, Y. Deng and J. Han, *Journal of Organic Chemistry*, 2018, **83**, 4674-4680.
24. H. E. Colley, M. Muthana, S. J. Danson, L. V. Jackson, M. L. Brett, J. Harrison, S. F. Coole, D. P. Mason, L. R. Jennings, M. Wong, V. Tulasi, D. Norman, P. M. Lockey, L. Williams, A. G. Dossetter, E. J. Griffen and M. J. Thompson, *Journal of Medicinal Chemistry*, 2015, **58**, 9309-9333.
25. O. Jentzer, P. Vanelle, M. P. Crozet, J. Maldonado and M. Barreau, *European Journal of Medicinal Chemistry*, 1991, **26**, 687-697.
26. J. L. Romera, J. M. Cid and A. A. Trabanco, *Tetrahedron Letters*, 2004, **45**, 8797-8800.
27. L. Tan, Z. Zhang, D. Gao, J. Luo, Z.-C. Tu, Z. Li, L. Peng, X. Ren and K. Ding, *Journal of Medicinal Chemistry*, 2016, **59**, 6807-6825.
28. A. F. Burchat, D. J. Calderwood, G. C. Hirst, N. J. Holman, D. N. Johnston, R. Munschauer, P. Rafferty and G. B. Tometzki, *Bioorganic & Medicinal Chemistry Letters*, 2000, **10**, 2171-2174.
29. P. Innocenti, H. L. Woodward, S. Solanki, S. Naud, I. M. Westwood, N. Cronin, A. Hayes, J. Roberts, A. T. Henley, R. Baker, A. Faisal, G. W.-Y. Mak, G. Box, M. Valenti, A. De Haven Brandon, L. O'Fee, H. Saville,

Chapter 4 – Lead Optimisation of Pyrrolopyrimidine-derived EGFR Inhibitors

- J. Schmitt, B. Matijssen, R. Burke, R. L. M. van Montfort, F. I. Raynaud, S. A. Eccles, S. Linardopoulos, J. Blagg and S. Hoelder, *Journal of Medicinal Chemistry*, 2016, **59**, 3671-3688.
30. W. Kabsch, *Journal of Applied Crystallography*, 1993, **26**, 795-800.
31. R. Read, *Acta Crystallographica Section D*, 2001, **57**, 1373-1382.
32. P. Emsley and K. Cowtan, *Acta Crystallographica Section D*, 2004, **60**, 2126-2132.
33. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallographica Section D*, 2010, **66**, 213-221.
34. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *Journal of Applied Crystallography*, 1993, **26**, 283-291.

Chapter 5

1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Abstract

In this chapter, our research endeavours departed from the traditional use of reversible inhibitors and acrylamide electrophiles which target cysteine residues within the kinase. To effectively combat the emergence of the C797S resistance mutation, we aspired to re-establish covalent inhibition in this mutant variant by targeting the EGFR catalytic lysine residue. The regulatory role of the catalytic lysine in kinase activation is discussed, followed by a review of several successful instances of lysine-targeted covalent inhibition. Emulating the successes in the literature, we devised a strategy employing 1,4-dicarbonyl electrophiles to potentially undergo Paal-Knorr pyrrole formation with the catalytic lysine, thereby covalently modifying the enzyme. This included use of 1,4-diketone, -dicarboxylic, -diester and -dialdehyde warheads attached to a suitably designed quinazoline driving group. Of these, eight final compounds were successfully synthesised, while several synthetic approaches to afford the 1,4-dialdehyde proved ineffective. Subsequent biochemical and cellular evaluation unfortunately yielded mostly poor results and covalent mass spectrometry experiments exhibited a reversible binding mode for all compounds. However, several unexpected outcomes that provide a platform for further development and future considerations to improve and enhance this class of inhibitor are disclosed.

5.1 Introduction

In the previous chapter, we utilised the traditional approach of reversible and irreversible inhibition, using the well-established acrylamide moiety as Michael acceptor to target the Cys797 residue for irreversible kinase inhibition. This conventional method of covalent modification rendered effective results against the L858R/T790M double mutant for which it was initially developed. The compounds exhibited exceptional inhibition of the double mutant biochemically and retained their efficacy in a cellular setting. However, as seen with the 3rd generation inhibitors osimertinib and olmutinib, a near complete loss of activity was witnessed against the clinically relevant C797S triple mutant, highlighting the ineffectiveness of this therapeutic strategy in targeting this mutant. The designed reversible inhibitors demonstrated impressive biochemical activity against the C797S mutant, with the most potent agent exhibiting a 75-fold increase over osimertinib. This advance in activity over osimertinib was preserved in the evaluation against the PC9/T790M/C797S cell line but was notably diminished to a mere 2-fold improvement.

A valuable lesson was instilled over a decade ago with the development and ultimate failure of 2nd generation inhibitors attempting to combat the emergence of the T790M resistance mutation. Trying to improve on existing inhibitors, as exemplified by the 2nd generation inhibitors derived from gefitinib and erlotinib, proved ineffective in quelling the advance of acquired drug resistance at the time. It was only through the combined discovery of novel chemical structures, and implementation of the then experimental acrylamide moiety for covalent modification, that 3rd generation inhibitors were able to effectively overcome the T790M mutation. The C797S triple mutant, which replaces the nucleophilic Cys797 with a serine residue, directly counteracts the fundamental functionality of this irreversible inhibition strategy. Consequently, the advent of this drug resistant mutant can be seen as a reflection of the situation researchers found themselves in at the inception of the EGFR-T790M mutant and will require the use of innovative therapeutic approaches to overcome.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Therefore, in the following chapters, our research thrust shifts from using the traditional irreversible therapeutic system of targeting cysteine residues with an electrophilic Michael acceptor. Instead, we endeavoured to design and synthesise inhibitors containing novel electrophiles which could potentially covalently modify relatively unexplored amino acids within the EGFR active site. We believe that the reinstatement of irreversible inhibition by these molecules could restore their cellular efficacy against the C797S triple mutant, a feat that no research group has accomplished up until this point. The body of research described in this chapter focusses on the use of 1,4-dicarbonyl electrophiles to potentially undergo Paal-Knorr pyrrole formation with the catalytic lysine 745 residue.

5.2 Targeting the EGFR Catalytic Lysine Residue

5.2.1 Catalytic Control in EGFR and Kinase Regulatory Mechanisms

The catalytic lysine residue is well-conserved in protein kinases and is involved in the phosphate transfer machinery of the substrate ATP molecule.¹ Using EGFR as an example in **Figure 5.1a**, we see that the catalytic Lys745 is an invariant residue of the $\beta 3$ strand, situated at the roof of the enzyme active site and in close proximity to helix αC of the N-lobe of the kinase. This residue resides in a similar position in virtually every ATP binding pocket of all human kinases.¹ As discussed in **Section 2.2.3** and **2.3.2**, protein kinases must undergo a conformation conversion to become catalytically active. One of the major modes by which kinase activation is regulated is by phosphorylation of their activation loop, which forms part of the activation segment and includes the DFG motif.² This mechanism induces the dynamic assembly of the regulatory spine, a well-documented part of the activation process and hallmark signature of an active kinase, and is accompanied by structural and conformational changes in the kinase.^{2, 3}

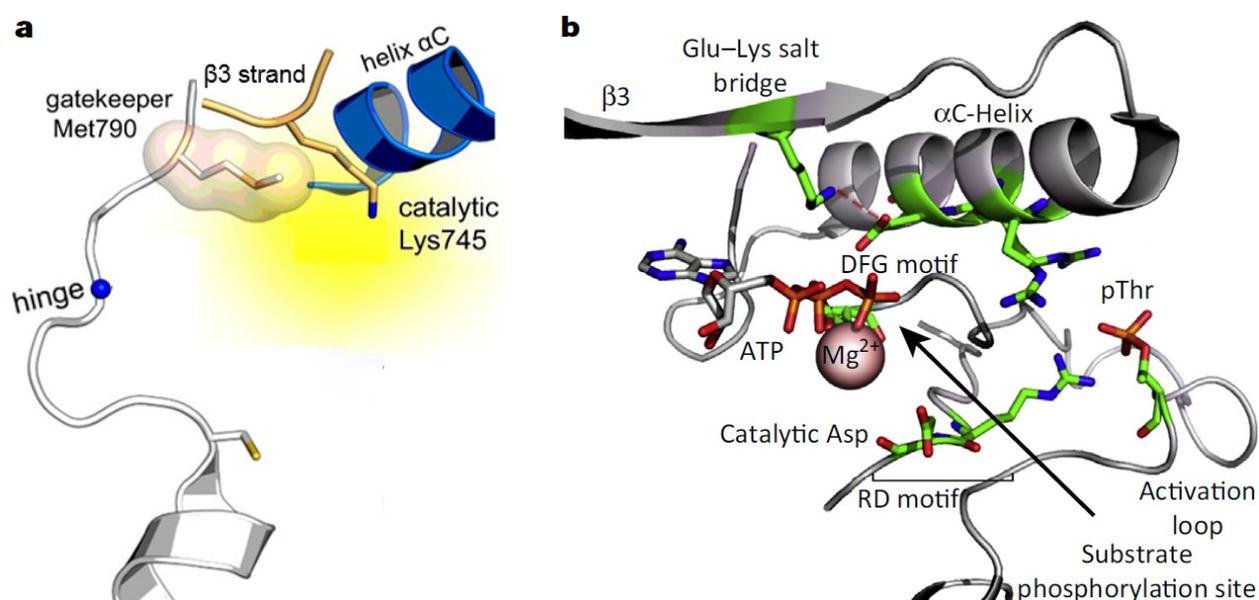


Figure 5.1: a) Crystal structure of the EGFR 696-1022/T790M mutant, highlighting the catalytic lys745 residue in yellow (PDB: 5GTZ)⁴ and b) The phosphorylated and active p38 α MAP kinase, with the activation-loop forming interactions with multiple regions, including the important Glu-Lys salt bridge (PDB: 3PY3). Figure redrawn and reproduced from a) Grabe et al and b) Beenstock et al.²

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Following activation-loop phosphorylation, a number of functional elements required for the activity of the protein kinase are aligned, the results of which can be seen in **Figure 5.1b**.² The activation loop, which is mostly disorganised in the inactive state, is stabilised and imposes a conformational change on the entire protein which promotes activity. The DFG motif undergoes a conversion to the DFG-in orientation, reorienting the regulatory spine and inducing rotation of the helix α C to α C-in.² It is this rotation that places the helix α C glutamic acid in a favourable position for the formation of the Glu-Lys salt bridge with the catalytic lysine (**Figure 5.1b**).² In this conformation, the lysine is able to interact with the α - and β -phosphates of ATP, stabilising it and facilitating transfer of the phosphate group within the substrate phosphorylation site.⁵ Therefore, the role of the catalytic lysine residue and formation of this salt bridge is an essential component for ATP binding and kinase activity.

The importance of the catalytic lysine residue is illustrated in examples of its mutation generating inactive or “kinase-dead” forms of the enzyme. In two separate studies, the catalytic lysine residue was mutated to a corresponding arginine and histidine residue in the MAP kinase ERK2 and cAMP-dependant kinase respectively.^{6,7} In the former, the mutation of the lysine residue created a non-productive binding mode for ATP and resulted in a dramatically lowered catalytic activity compared to the wild-type ERK2 enzyme, suggesting its importance in orienting ATP and enabling catalysis.⁶ Similar results were obtained in the latter investigation, with the mutant variant displaying a lower stability than the wild-type catalytic subunit.⁷

Lastly, crystal structure studies of **lapatinib**, an FDA approved dual EGFR/HER2 inhibitor shown in **Figure 5.2a**, revealed a previously unseen inactive-like conformation of the EGFR enzyme. In this binding mode, the helix α C is displaced to the α C-out conformation, while the DFG motif remains in DFG-in.⁸ Following the helix α C movement, the conserved Glu-Lys salt bridge interaction is broken, inactivating the enzyme.

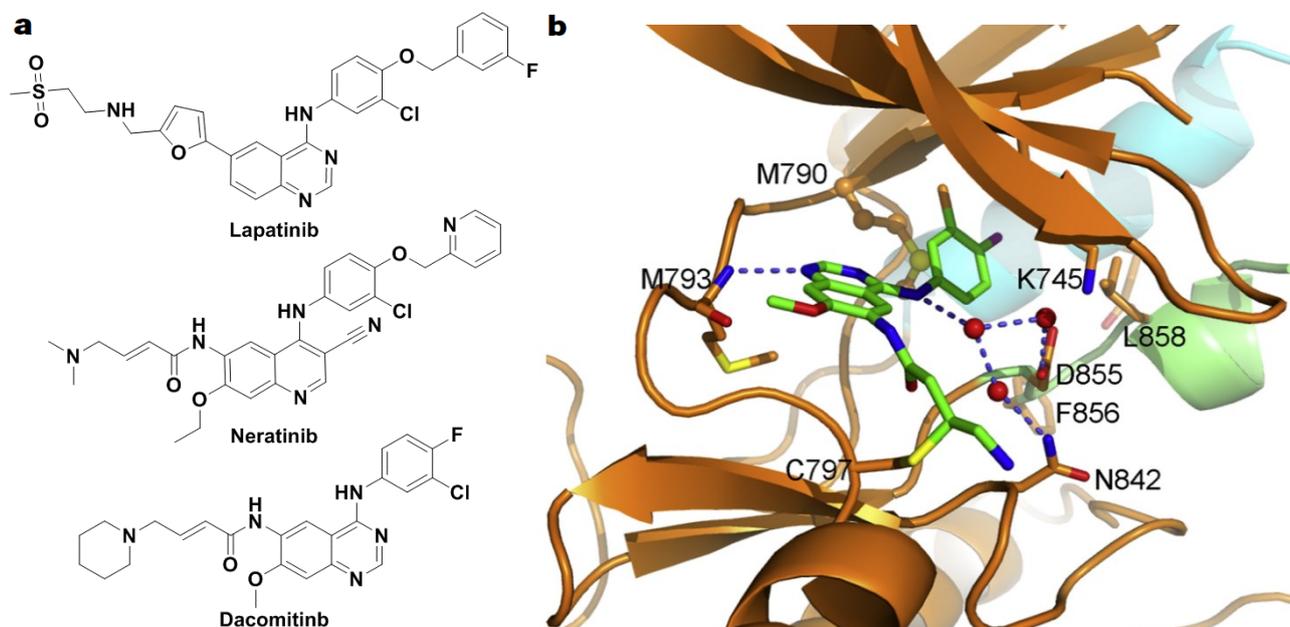


Figure 5.2: a) Structures of lapatinib, neratinib and dacomitinib and b) Crystal structure of dacomitinib, with the activation loop highlighted in lime-green, helix α C in light-blue and catalytic lysine labelled K745 with the Glu-Lys salt bridge absent. Figure reproduced from b) Gajiwala et al.⁹

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Similar findings were obtained for both **neratinib** and **dacomitinib**, shown in **Figure 5.2a**, with both inhibitors displaying the DFG-in α C-out binding mode.^{9, 10} A crystal structure of dacomitinib, bound in this conformation to the EGFR-T790M mutant, is shown in **Figure 5.2b**.⁹ In this binding mode we see the helix α C (**Figure 5.2b**, light-blue) clearly in the out-conformation and a fair distance from the catalytic lysine K745. As the glutamic acid residue of the salt bridge is located on the helix α C, and owing to this distance, the salt bridge interaction is broken in this orientation. This highlights the importance of the salt bridge, amongst other structural features, in maintaining activation of the protein kinase. However, with the increasing knowledge of the role of the DFG-in and helix α C-out form of the enzyme, recent protein kinase inhibitors have been designed to target this specific inactive binding.¹¹

5.2.2 Lysine-trapping Covalent Inhibitors

Whilst the vast majority of irreversible inhibitors to date have been designed to target a nucleophilic cysteine residue in close proximity to the ligand binding site, examples of lysine-targeting covalent inhibitors, although exceedingly rare, have been reported.¹² Originally, natural products were the only source of lysine-trapping covalent compounds. Isolated as an antibiotic produced by *Penicillium wortmanni* in 1957, the natural product **wortmannin**, shown in **Figure 5.3a**, has been shown to irreversibly inhibit the PI3K family of kinases. Depending on the isoform, these oncologically relevant protein kinases are covalently modified at the catalytic Lys802 or Lys833 residue within the ATP binding site.¹³ Through protein crystal structure studies, the mechanism of action of wortmannin was revealed to involve nucleophilic attack of the activated furan ring by the catalytic lysine residue, leading to its ring opening and finally the covalently bound enamine species shown in **Figure 5.3a**.¹⁴ The crystal structure of wortmannin covalently bound to PI3K, confirming the exact nature of the binding mode and the opened furan ring, can be seen in **Figure 5.3b**.¹² Wortmannin analogues have been successfully developed with the aim of increasing selectivity and reducing off-target toxicity, with PX-866 having progressed to a phase II clinical trial for treatment against NSCLC and glioblastoma.¹⁵

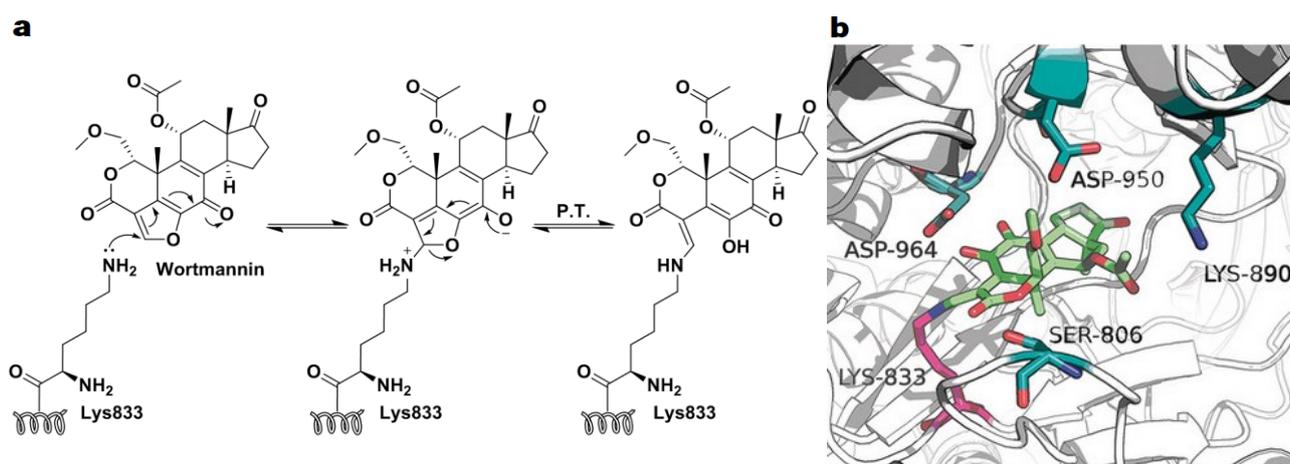


Figure 5.3: a) The proposed mechanism of action for wortmannin reacting with the catalytic Lys833 of PI3K and b) Crystal structure of PI3K covalently bound to wortmannin (green) through the catalytic Lys833 (magenta) (PDB: 1E7U). Figure reproduced from b) Pettinger, Jones and Cheeseman.^{12, 14}

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

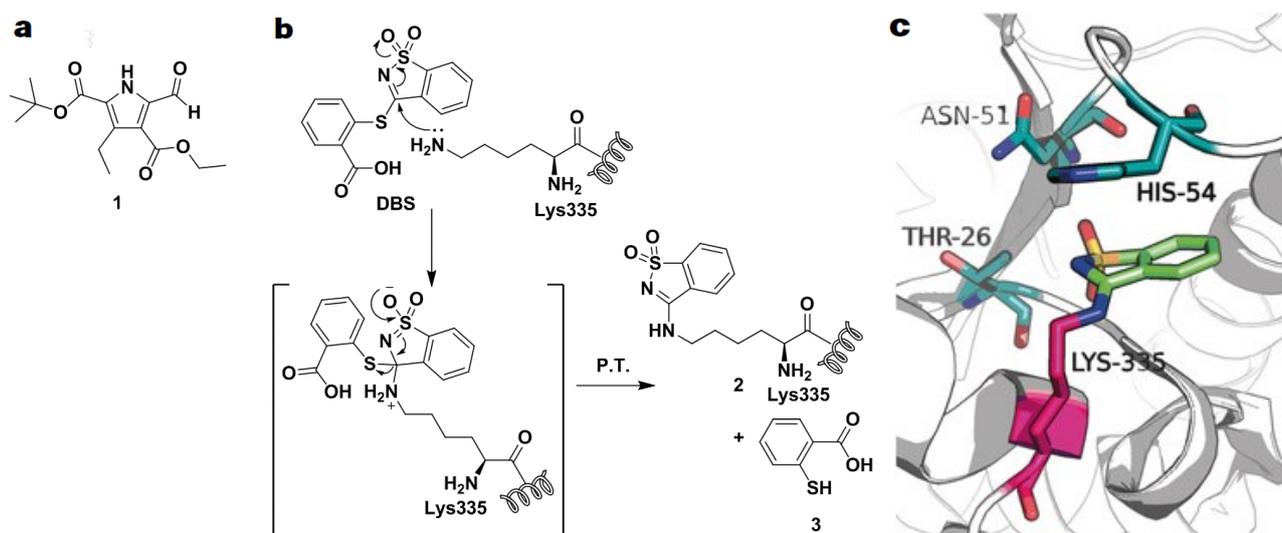


Figure 5.4: a) Structure of IGF-1R inhibitor **1**, b) Structure and mechanism of irreversible inhibition of DBS, with covalent adduct **2** and by-product **3** and c) Crystal structure of **2** (green) covalently bound to Lys335 (magenta) of the LmPYK enzyme (PDB: 3SRK). Figure reproduced from c) Pettinger, Jones and Cheeseman.^{12, 16, 17}

Bell and co-workers developed a range of pyrrole-5-carboxaldehyde-derived, type 1 insulin-like growth factor receptor (IGF-1R) inhibitors which showed modest activity and good selectivity over a variety of kinases.¹⁶ These aldehyde-based inhibitors, with compound **1** as example illustrated in **Figure 5.4a** above, underwent reversible covalent imine bond formation between the aldehyde and Lys1003 within the ATP-binding pocket. To understand the mechanism of action, the aldimine adduct was reduced within IGF-1R using sodium borohydride *in situ*. This resulted in irreversible inactivation of the kinase, allowing for subsequent protein crystallography and confirming the formation of a reversible covalent bond. Compound **1** demonstrated an IC_{50} value of $0.49 \mu\text{M}$ against IGF-1R.

The discovery of irreversible inhibitors using high-throughput screening (HTS) is a challenging endeavour as the inherent reactivity of the electrophile bearing small molecules typically excludes them from screen collections.¹⁸ However, Morgan et al. were able to discover a new class of irreversible inhibitor during a HTS targeting *Leishmania Mexicana* pyruvate kinase (LmPYK).¹⁷ Following limited development of the hit compound, the saccharin derived thiol-linked analogue **DBS**, pictured in **Figure 5.4b**, was found to demonstrate good biochemical potency and inhibitory activity in a time- and dose-dependent manner. Crystallisation of DBS with LmPYK revealed the saccharin moiety undergoing covalent bond formation by nucleophilic attack from the active site Lys335 residue, forming the covalent adduct **2** and leaving group **3** shown in **Figure 5.4b**.¹⁷ The complex of **2** can be seen in **Figure 5.4b**.¹²

The research group of Shokat were able to develop a chemical method that allowed for covalent cross-linking of a substrate of interest to protein kinase B (AKT).¹⁹ The novel mechanism-based cross-linker **4**, shown in **Figure 5.5a**, was derived from the general covalent linker 5'-fluorosulfonylbenzoyl adenosine (FSBA) and contained an *o*-phthaldialdehyde moiety. Use of this dialdehyde catered for the requirement of a bifunctional cross-linker and specifically targeted the invariant catalytic lysine within the ATP binding site of the kinase. Following formation of the reversible imine adduct **5** (**Figure 5.5a**), nucleophilic attack of the cysteine-containing substrate and subsequent hydrolysis results in the formation of the stable isoindole linked covalent complex **6** (**Figure 5.5a**). This covalent cross-linker proved to be robust and was used in the identification of various kinase-substrate pairs.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

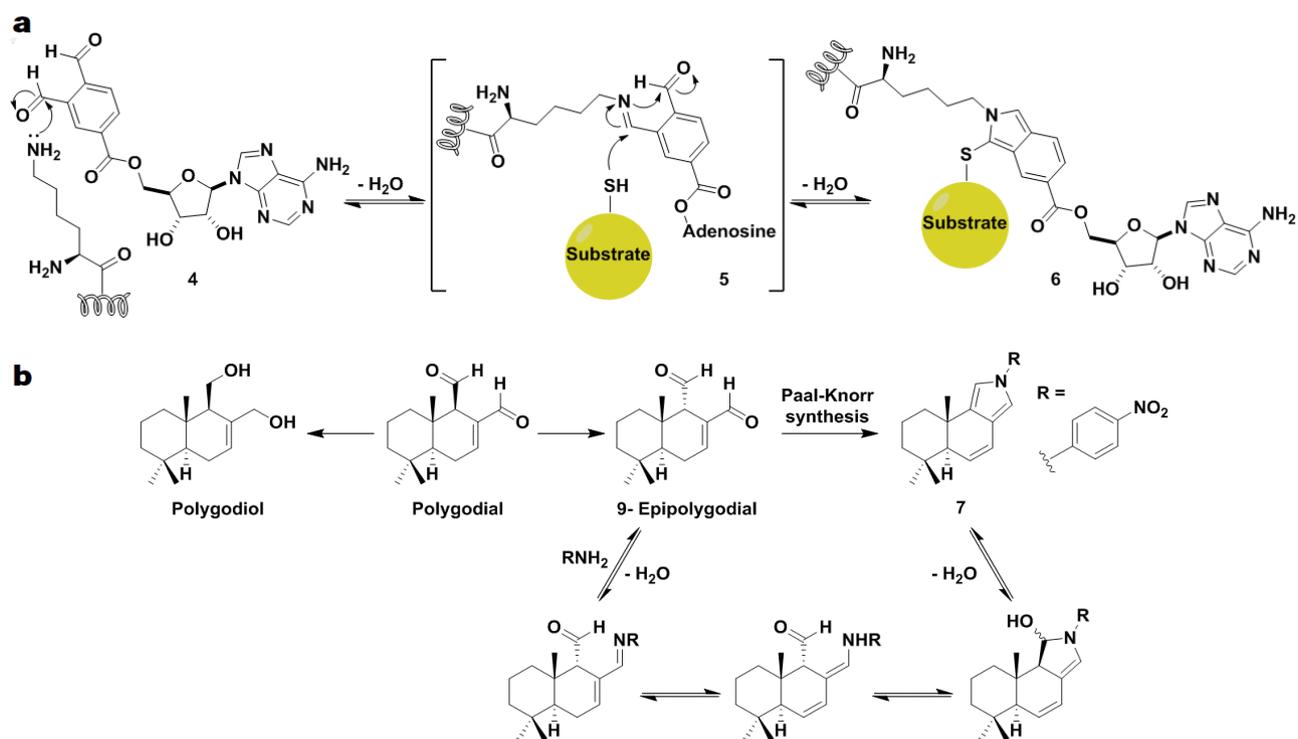


Figure 5.5: a) Representation of the covalent cross-linker **4** and cross-linking reaction furnishing the substrate-kinase pair **6** through intermediate **5** and b) The natural product polygodial and its derivatives, polygodiol and 9-epipolygodial. Reversible synthesis of the stable Paal-Knorr pyrrole is also shown.^{19, 20}

The natural product **polygodial**, shown in **Figure 5.5b**, is a terpenoid dialdehyde isolated from *Polygonum hydropiper* L. and is a known agonist of transient receptor potential vanilloid (TRPV1).²¹ Kornienko and co-workers, our collaborators from Texas State University, prepared and investigated a series of polygodial analogues for their TRPV1-agonist and anticancer properties.²⁰ Through these experiments, the group was able to identify and isolate the epimer **9-epipolygodial** (**Figure 5.5b**), which displayed potent antiproliferative activity, exceeding that of the parent compound, against apoptosis- and multidrug-resistant cancer cells. Attempting to understand the mechanism of action, the diol derivative **polygodiol** (**Figure 5.5b**) was synthesised, leading to a complete loss of anticancer activity and highlighting the importance of the dialdehyde functionality. The group hypothesised that the dialdehyde moiety could be undergoing a modified Paal-Knorr condensation with a lysine residue on the target protein. To support this intriguing claim, a chemical feasibility study was successfully undertaken, resulting in the synthesis of the stable polygodial pyrrole derivative **7**, through the proposed pathway (**Figure 5.5b**). Exploration and use of this novel mechanism of covalent modification could be thus beneficial in the development of new anticancer therapeutic agents.

The examples discussed above highlight a few applications of lysine-targeting electrophiles. Further examples are implemented for use in synthetic affinity labels, activity-based protein profiling and inhibitors of other enzymatic targets.¹² The recent developments in rational targeting of nucleophilic lysine residues demonstrates its potential for future drug discovery. This provided the inspiration for the body of research described in the rest of this chapter.

5.3 Objectives and Target Synthetic Library

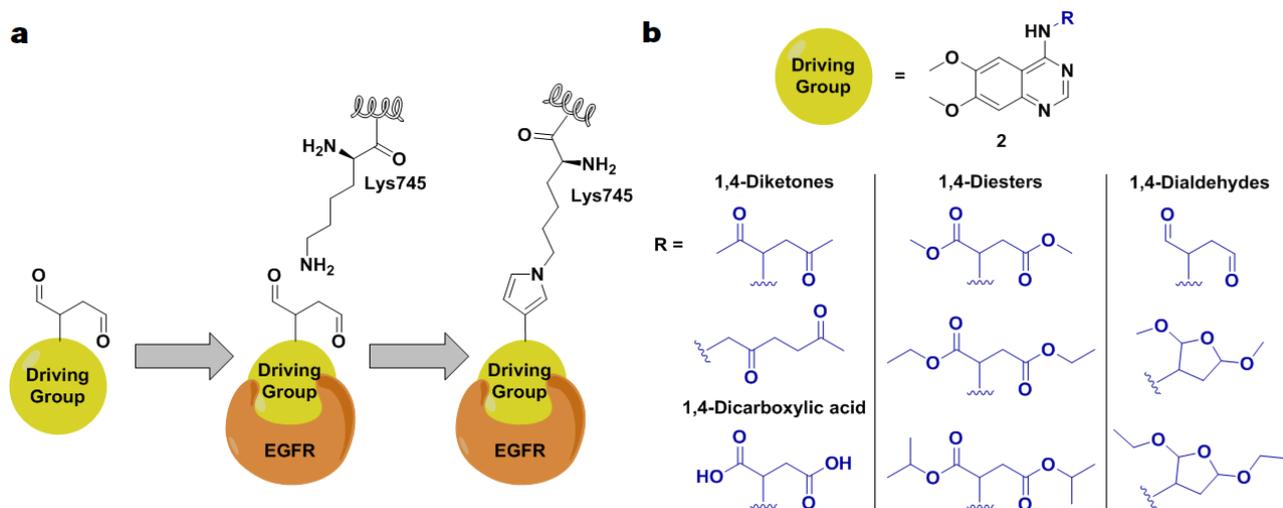


Figure 5.6: a) Use of a 1,4-dicarbonyl electrophile to potentially undergo Paal-Knorr pyrrole synthesis with the EGFR catalytic Lys745 residue and b) Illustration of the selected heterocyclic driving group and various 1,4-dicarbonyl electrophiles constituting the prospective synthetic library.

At the time of undertaking this research endeavour, the EGFR catalytic Lys745 residue had not been covalently modified by a corresponding irreversible inhibitor. Owing to the basicity of the ϵ -amino group of a typical surface lysine (pK_a of 10.4) and physiological pH being at 7.4, approximately 99.9% of lysine residues within the human body will be presented as the protonated species and thus unavailable as nucleophiles.¹² Consequently, we were curious as to whether the catalytic Lys745 of EGFR, situated relatively deep within the ATP binding pocket, could act as a nucleophile and undergo covalent bond formation. If indeed possible, we were intrigued by the potential therapeutic aspects of this mechanism of irreversible inhibition. As the catalytic lysine plays a significant role in the activation of the kinase and phosphate transfer of ATP, covalent modification of this residue would negate its ability to undergo salt bridge formation, leaving the kinase permanently inactivated. As most oncogenic kinases are in a permanently activated state, the prospect of being able to “switch off” these mutant variants of the enzyme by targeting a component of its machinery seemed advantageous.

Following suit from the examples discussed in the previous section and the work undertaken by Professor Kornienko with the dialdehyde containing polygodial, we devised a strategy to covalently modify the catalytic Lys745 by using a 1,4-dicarbonyl electrophile. As illustrated in **Figure 5.6a**, the warhead would be attached to a suitable heterocyclic driving group with high reversible binding affinity, able to form important interactions with the hinge region of the kinase. Following reversible binding, we envisaged the electrophile undergoing a Paal-Knorr condensation reaction with the catalytic Lys745 residue, resulting in the formation of a covalently bound pyrrole (**Figure 5.6a**).

As discussed in **Section 5.2.1**, the quinazoline-based inhibitors of lapatinib, neratinib, dacomitinib and gefitinib possess the ability to bind to both the traditional, active conformation of the kinase and the distinct DFG-in α C-out inactive conformation of the kinase. For the purpose of our investigation we found this feature desirable, as in the active conformation of the kinase covalent modification of the catalytic Lys745 could result in its potential deactivation, whilst in the inactive conformation the absence of the Glu-Lys salt bridge could potentially present a more nucleophilic lysine residue to undergo this covalent bond formation.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Furthermore, owing to their significant biological activity and frequent use in kinase inhibitors, a wealth of knowledge exists on quinazoline-based derivatives, providing us with a relatively simplified synthetic route towards these structures.²² For these reasons, we selected a quinazoline core as our heterocyclic driving group, which can be seen at the top of **Figure 5.6b**. Lastly, as this study would be focussed on a proof of concept approach, we decided to forego the conventional inclusion of a solubilising group and opted instead for the simplified disubstituted methoxy at the 6- and 7-positions on the aromatic ring (**Figure 5.6b**).

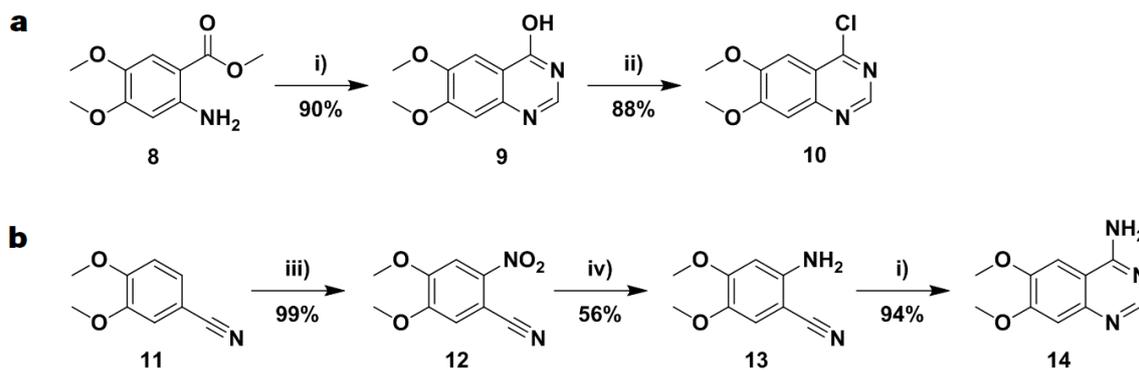
We elected to employ a variety of 1,4-dicarbonyl electrophiles to potentially undergo the Paal-Knorr pyrrole synthesis with the catalytic lysine residue. These included the use of 1,4-diketone, 1,4-dicarboxylic acid, 1,4-diester, 1,4-dialdehyde and masked 1,4-dialdehyde functionalities, as can be seen in the bottom right of **Figure 5.6b**. We theorised that use of different carbonyl groups with varying reactivities and steric bulk/leaving groups could provide insight into the rate at which the Paal-Knorr condensation occurs, tailored lipophilicity and stability of the formed covalent adducts. As the amine at the 4-position of most commercial quinazoline inhibitors forms a hydrogen bond with the solvent in the ATP binding site (**Figure 5.2b**), we deemed it important to preserve this interaction in our library of compounds. Additionally, the halogenated aniline fragments which are substituted at the 4-position in these clinical quinazoline inhibitors (**Figure 2.11** and **Figure 5.2**) bind in a close proximity to the catalytic lysine residue.⁹ We therefore chose to derivatise with our 1,4-dicarbonyl electrophiles at this 4-aminoquinazoline position, placing the warheads in a theoretically favourable and nearby position to undergo the Paal-Knorr condensation reaction with Lys745. With our objective of covalently modifying the EGFR catalytic lysine residue and our initial synthetic library in mind, we set about synthesising the potential irreversible inhibitors.

5.4 Library Synthesis

5.4.1 Heterocyclic Driving Group Synthesis

To facilitate the future incorporation of our 1,4-dicarbonyl electrophiles at the 4-position of the 6,7-dimethoxyquinazoline driving group, the heterocyclic scaffolds **10** and **14**, shown in **Scheme 5.1a**, were synthesized. For the synthesis of compound **10**, the first step required cyclization of the commercially available 4,5-dimethoxyanthranilic acid **8** with formamide. The optimal reaction temperature was found to be a maximum of 140 °C, with anything above this resulting in a marked decrease in yield. Additionally, the inclusion of ammonium formate over the use of formamide exclusively as solvent, effected a further improvement in the yield. When heated, ammonium formate undergoes elimination of water, releasing formamide. This provided the tautomeric compound **9** in an excellent yield of 90%. This was followed by conversion of the **9** into the corresponding 4-chloroquinazoline. Initial attempts using thionyl chloride and phosphorus oxychloride required harsh conditions of high temperature and excessive amounts of the chlorinating reagents, resulting in a lower yield and undesirable purification procedures. However, the addition of a catalytic amount of DMF prompted *in situ* formation of the reactive Vilsmeier-Haack reagent, allowing for lower temperatures and lesser equivalents of thionyl chloride to be used. Under these conditions, the first heterocyclic driving group scaffold **10** was therefore afforded in good yield.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745



Scheme 5.1: Synthesis of the heterocyclic driving group scaffolds **10** and **14**. Reagents and conditions: i) ammonium formate (1.5 equiv.), formamide, 140 °C, 8 h; ii) SOCl₂ (20 equiv.), DMF (cat.), 80 °C, 6 h; iii) HNO₃ (10 equiv.), Ac₂O, -10 °C – rt, 12 h; iv) Na₂S₂O₄ (8 equiv.), TBAB (30 mol%), CHCl₃/H₂O (1:1), rt, 16 h.²³

The outset of the synthesis of **14**, illustrated in **Scheme 5.1b** above, required the regioselective nitration of the commercially available 3,4-dimethoxybenzotrile **11**. Dropwise treatment of **11** with an excess of nitric acid in acetic anhydride, while maintaining a temperature below -5 °C, followed by precipitation in water and subsequent filtration, furnished the nitrated derivative **12** in near quantitative yield with no further purification required.²³ While the combined activating/deactivating effects of the ring's substituents indicate nitration at the 5-position, we believe the steric burden of a tri-substituted aromatic product outweighs these effects. Furthermore, the obtained ¹H NMR spectrum compared well with the reported literature values, negating the requirement for further verification.²³

Due to the presence of the benzonitrile moiety, we excluded the use of hydride-based reducing agents for the aromatic nitro reduction of compound **12**, thus avoiding the possible formation of the diamine product. Use of sodium dithionite, with tetrabutylammonium bromide (TBAB) as phase transfer catalyst in a mixture of chloroform and water, facilitated chemoselective reduction of the nitro functionality, affording compound **13**. Optimization efforts, including use of DCM and CCl₄ in different ratios, higher temperatures and varying equivalents of the reducing agent could not drive the reaction to completion. The reported literature yield of 56% was matched and unreacted starting material was iteratively recycled and used in subsequent reductions.²³ Lastly, ring closure of compound **13** with the previously established ammonium formate/formamide conditions supplied the 4-aminoquinazoline scaffold **14** in excellent yield.

Owing to the structural similarity of the driving group scaffolds **10** and **14**, only the ¹H NMR spectrum of compound **14** will be depicted and scrutinized. Full characterization of all compounds not analysed in this chapter, including IR, ¹H, ¹³C and HRMS spectra, can be accessed in the supplementary information (**Section 5.8**). Inspection of the ¹H NMR spectra in **Figure 5.7** indicated structural confirmation of compound **14**. In the upfield region, we saw two singlets at δ 3.86 and 3.89 ppm, integrating for 3H each and correlating to the 6- and 7-position methoxy groups (dark blue). Moving downfield, a characteristic broad peak of 2H was seen at δ 7.41 ppm, corresponding to the 4-amino moiety (orange). Lastly, we anticipated the 2-position proton (red) to have the largest downfield shift, owing to the significant deshielding effects of the neighbouring nitrogen atoms in the quinazoline ring. This singlet appeared at δ 8.26 ppm, with the remaining aromatic protons (green) assigned to the peaks at δ 7.07 and 7.57 ppm. With the heterocyclic driving group scaffolds in hand, we next set about synthesising our 1,4-dicarbonyl electrophiles.

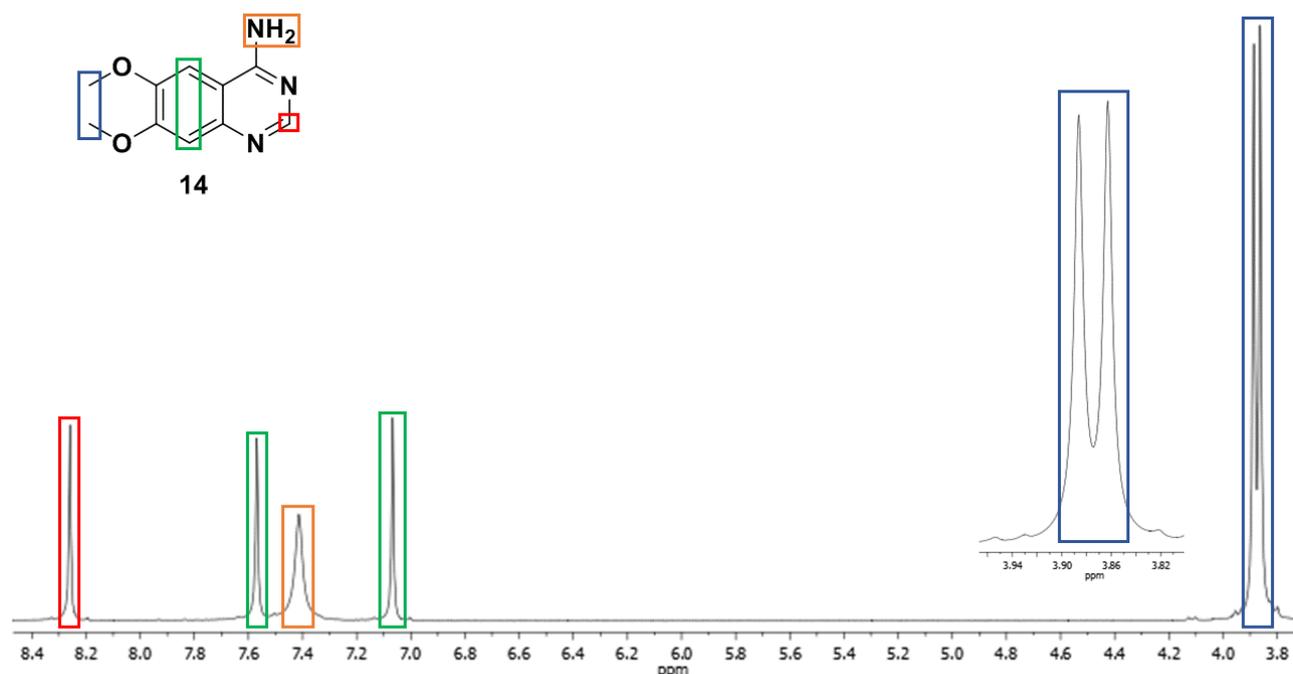
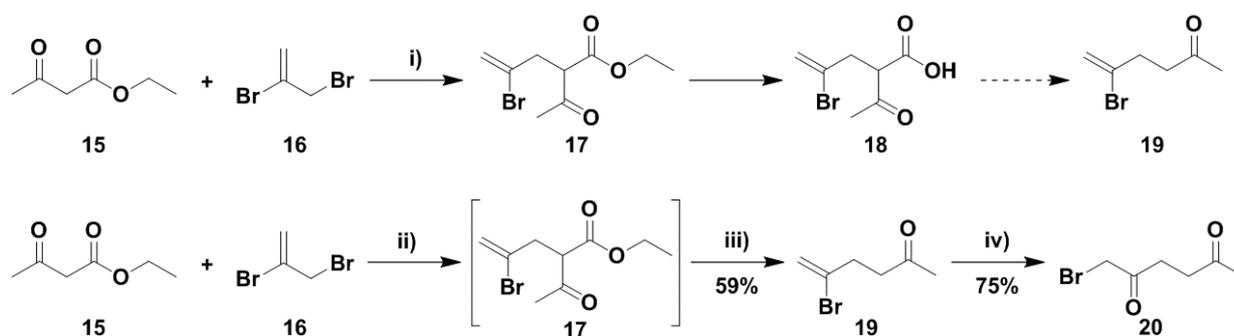


Figure 5.7: ^1H NMR Spectra of the heterocyclic driving group compound **14**.

5.4.2 1,4-Diketone Compound Synthesis

For the synthesis of the 1,4-diketone-containing compound, we envisaged coupling of the synthesised scaffold **14** to the α -bromoketone electrophile **20**, the synthesis of which can be seen in **Scheme 5.2**. Synthesis of electrophile **20** began with $\text{S}_{\text{N}}2$ reaction between the commercially available ethyl acetoacetate and 2,3-dibromopropene, compounds **15** and **16** respectively. Following vacuum distillation purification of both reagents prior to reaction, use of *in situ* generated sodium ethoxide as base furnished the intermediate enolate, allowing for nucleophilic attack on the electrophilic **16**, affording compound **17** in good yield. Subsequent efforts to produce compound **19** via hydrolysis and decarboxylation of compound **17** were met with failure, repeatedly leading to isolation of the carboxylic acid **18**, represented by a broad peak near δ 12.3 ppm in the ^1H NMR spectra. This included the use of high temperatures, varying concentrations of acid and base and Krapcho decarboxylation conditions using LiCl and DMSO, all to no avail.



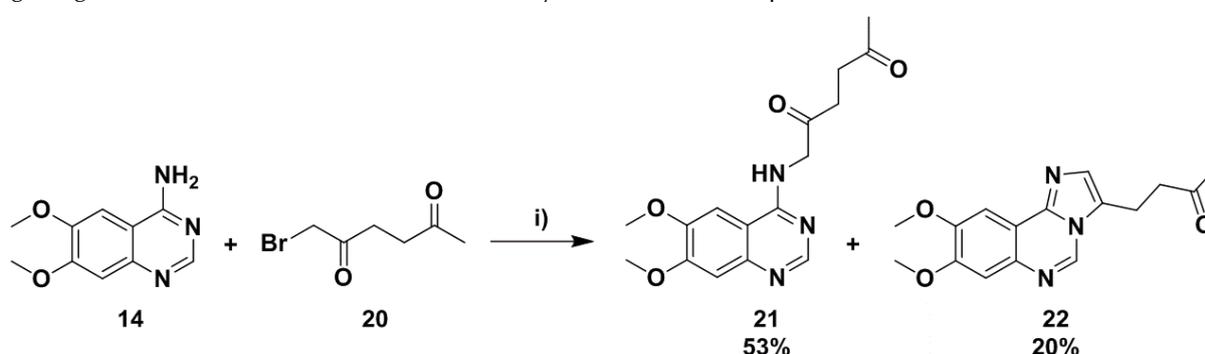
Scheme 5.2: Synthesis of the 1,4-diketone electrophile **18**. Reagents and conditions: i) NaOEt (generated *in situ*) (1 equiv.), EtOH/Et₂O (2:1), -10°C – rt, 24 h; ii) NaOEt (generated *in situ*) (1 equiv.), EtOH, 80°C , 1 h; iii) 10% NaOH, rt – 80°C , 2 h then conc. HCl, rt – 80°C , 2 h; iv) NBS (1.3 equiv.), HBr (cat.), MeCN:H₂O (4:1), rt, 4 h.^{24, 25}

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

We therefore returned to the initial S_N2 reaction and, following the formation of intermediate **17**, were able to induce hydrolysis and decarboxylation using NaOH and HCl at 80 °C in a one-pot synthesis.²⁴ This afforded the desired compound **19** in an acceptable yield of 59% over the three-step reaction sequence. The final step in the synthesis of electrophile **20** involved the oxidative hydrolysis of vinyl bromide **19** using NBS and a catalytic amount of hydrobromic acid in aqueous acetonitrile, a procedure developed by Morton and Leanna.²⁵ Addition of these reagents generates a bromine-hypobromous acid equilibrium, the latter species being responsible for the oxidative hydrolysis. This approach furnished the α -bromoketone electrophile **20** in good yield, setting up for subsequent coupling with scaffold **14**. It is worth noting that monobromination of 2,5-hexanedione was ruled out as a potential synthetic route towards **20**. Competing sites within the molecule gives rise to a lack of control in regioselectivity and over-bromination, with resultant product separation of the mono- and dibrominated products near impossible.

Reaction of scaffold **14** with electrophile **20** in ethanol while heating under reflux, as outlined in **Scheme 5.3** below, led to the formation of the target compound **21** and the ring-closed imidazoquinazoline equivalent **22** as the major and minor products respectively. This result is comparable to the work undertaken by the Cottam research group, where similar reaction conditions led to production of a series of small heterocyclic compounds including the benzothiazoles, benzimidazoles and quinazolines, together with their corresponding ring-closed imidazo-counterparts.²⁶ As **22** contains a suitable ketone functionality, we proposed that this electrophile could potentially undergo reversible imine formation with the catalytic Lys745, in similar manner to compound **1** (**Figure 5.4a**). Consequently, we decided to evaluate both compounds **21** and **22** against EGFR and its mutant variants.

The successful synthesis of both **21** and **22** could be confirmed through the comparative analysis of their respective ^1H NMR spectra, as illustrated in **Figure 5.8** on the following page. Examination of the upfield region in both spectra revealed the presence of several shared structural elements. These included the terminal ketone methyl peak (yellow) between δ 2.10 and 2.15 ppm, the signals of the 1,4-diketone ethylene bridge (light blue) between δ 2.81 and 2.91 ppm and the aromatic dimethoxy groups (dark blue) at approximately δ 4.00 ppm, with appropriate peak integration. Moving to the aromatic region, we could see the signal corresponding to the 2-position quinazoline proton (red) furthest downfield in both spectra. For compound **21** and **22** we found two more aromatic peaks (green) correlating to the remaining quinazoline protons. However, compound **22** contained an extra peak (green) integrating for 1H, which we attributed to the newly formed imidazole proton.



Scheme 5.3: Synthesis of the final 1,4-diketone compound **21** and the imidazoquinazoline ketone compound **22**. Reagents and conditions: i) EtOH, 80 °C, 6 h.²⁶

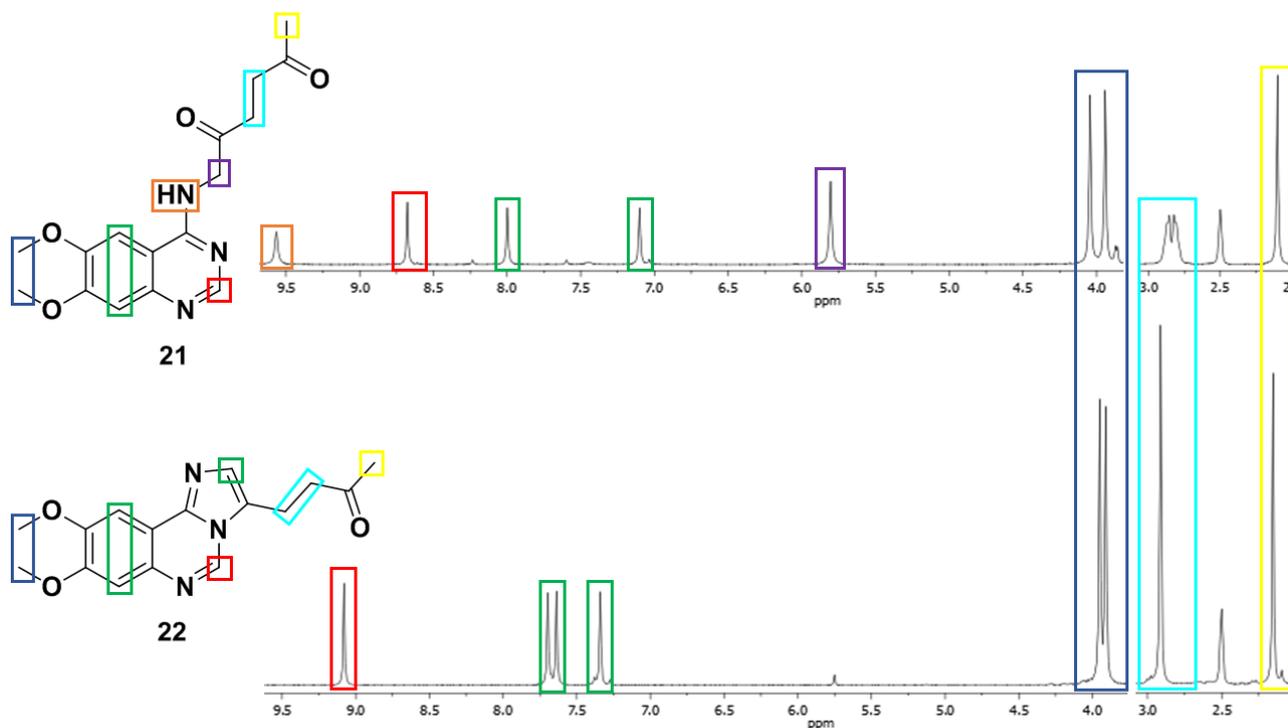


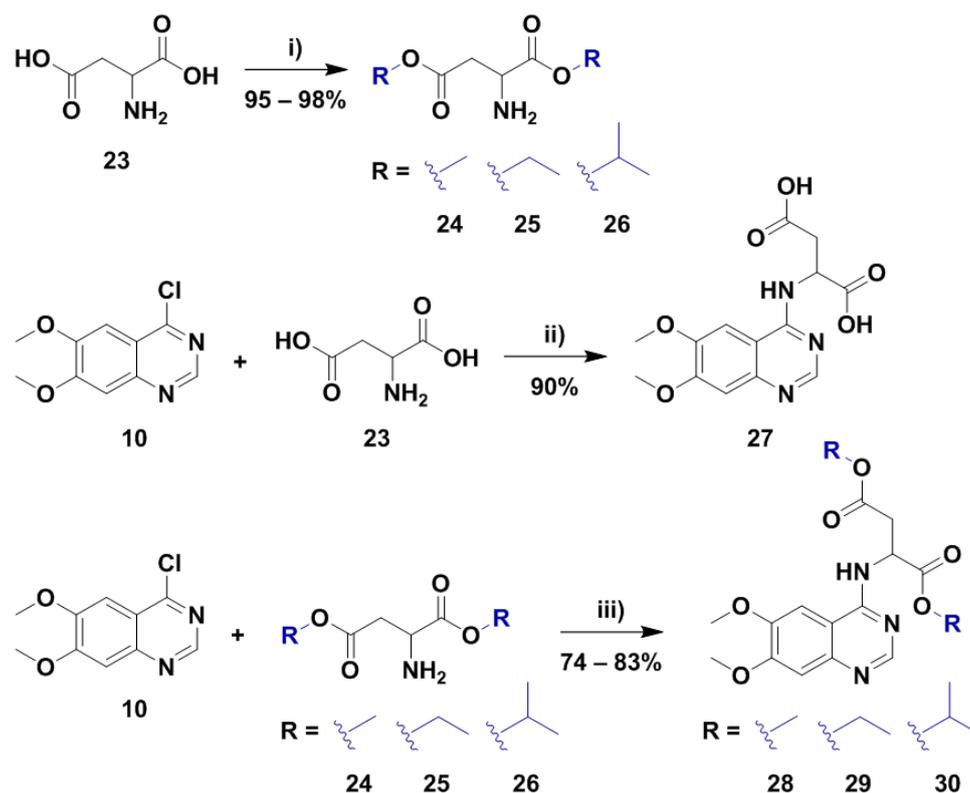
Figure 5.8: Comparison of the ^1H NMR spectra of final compounds **21** and **22**.

Lastly, a singlet belonging to the linker protons (purple) between the 1,4-diketone chain and the 4-aminoquinazoline ring in compound **21** was visibly present at δ 5.81 ppm and evidently absent in the spectra of compound **22**, further substantiating formation of the ring-closed product. Similarly, the 4-aminoquinazoline proton (orange) was found furthest downfield at δ 9.57 ppm in the spectra of compound **21** and was perceptibly lacking in the spectra of compound **22**. The absence of these signals in compound **22** signified their participation in the formation of the imidazoquinazoline. While not analysed, the ^{13}C spectra and HRMS results provided further confirmation for the formation of these two products and can be found in the supplementary information (**Section 5.8**). This afforded two ketone-based final compounds for biochemical and cellular evaluation and prompted us to move forward in the synthesis of the remaining final compounds.

5.4.3 1,4-Dicarboxylic Acid and 1,4-Diester Compound Synthesis

Synthesis of the 1,4-dicarboxylic acid and 1,4-diester-containing final compounds required the use of DL-aspartic acid as a building block. As outlined in **Scheme 5.4** on the following page, we envisaged reaction at the electrophilic 4-position of the quinazoline driving group **10** with the naked amino acid **23** and the esterified counterpart **24** – **26** acting as nucleophiles to furnish final compounds **27** – **30**. Accordingly, aspartic esters **24** – **26** were synthesised by treatment of **23**, in a solution of either methanol, ethanol or isopropanol, with thionyl chloride at low temperature to generate the aspartic acyl chloride intermediate. The solution was subsequently heated under reflux for 6 hours, allowing for nucleophilic attack by the respective alcohol, followed by distillation of excess thionyl chloride and trituration with diethyl ether. This furnished the aspartic ester derivatives **24** – **26** as the hydrochloride salts in excellent yields with no further purification required (**Scheme 5.4**).

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745



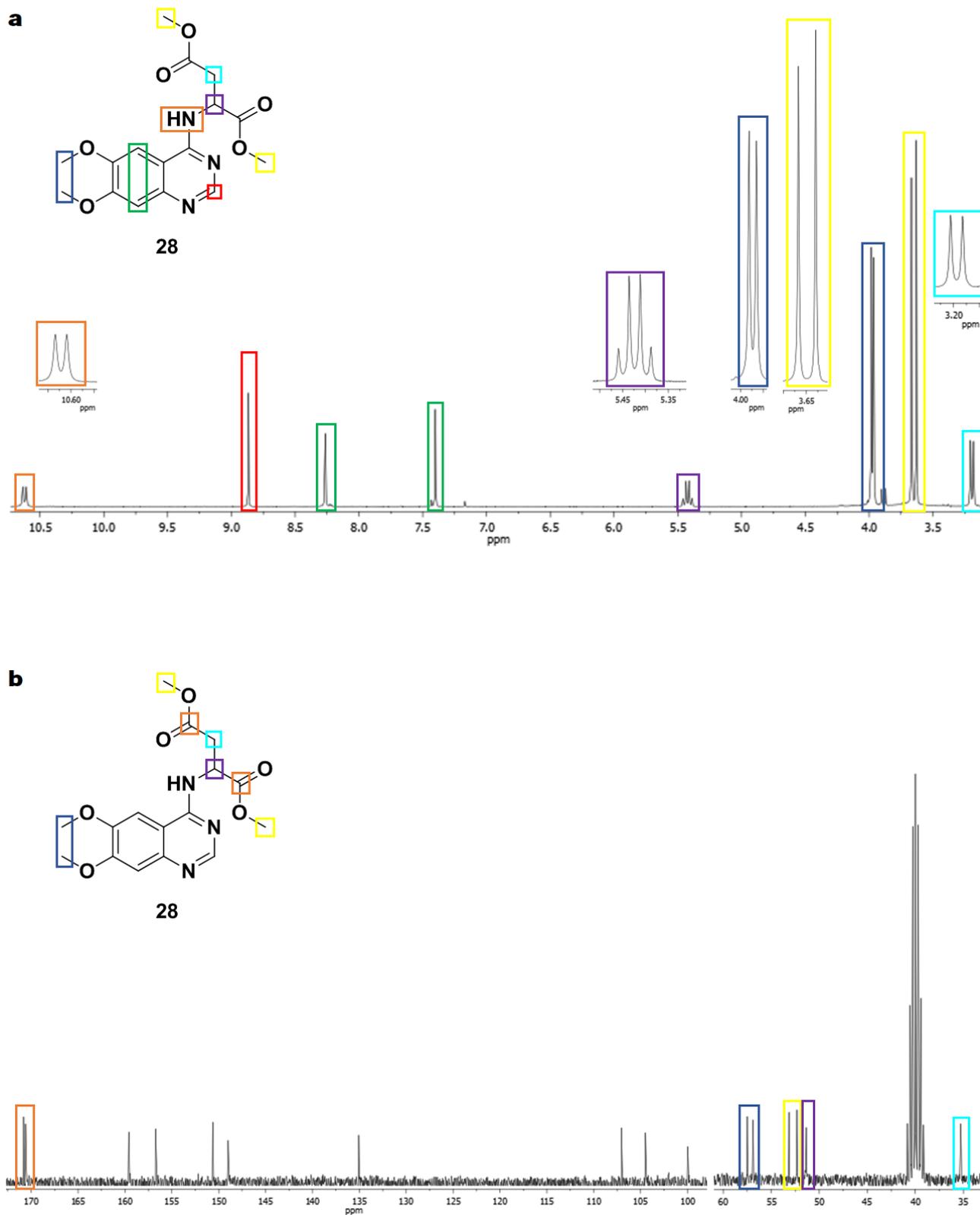
Scheme 5.4: Synthesis of the final 1,4-dicarboxylic acid compound **27** and the 1,4-ester containing compounds **28** – **30**.

Reagents and conditions: i) SOCl_2 (1.4 equiv.), respective alcohol, -10°C – reflux, 6 h; ii) DL-aspartic acid (0.83 equiv.), Et_3N (4 equiv.), $\text{DMSO}:\text{H}_2\text{O}$ (1:1), 80°C , 72 h; iii) aspartic ester (1.1 equiv.), 4 M HCl in dioxane (cat.), MeCN, 80°C , 24 h.

Formation of the 1,4-dicarboxylic acid-containing final compound **27** proved more challenging than anticipated. Solvent choice was problematic owing to the polarity of the reagents and formed product and the inability to use solvents such as methanol or ethanol which could act in a nucleophilic manner. Optimization parameters of the reaction included the interchanging use of acids or bases in various solvents at different temperatures. It was found that use of **23** as the limiting reagent aided in the subsequent purification and a 1:1 mixture of $\text{DMSO}/\text{H}_2\text{O}$ as solvent provided sufficient solubility and miscibility for both reactants. Employing the optimised conditions, coupled with the use of excess Et_3N as base at 80°C , furnished the final compound **27** in an excellent yield of 90%.

Due to the improved solubility of the aspartic acid derivatives **24** – **26**, optimization of their reaction with driving group **10** was uncomplicated. Under basic conditions the reaction proceeded sluggishly and did not reach completion, whereas the use of a catalytic amount of 4 M HCl in dioxane required somewhat shorter reaction times of up to 24 hours. We believe this increase in reaction rate stems from $N1$ protonation of the quinazoline heterocycle, with the subsequent resonance structure increasing the electrophilicity of the 4-position, thereby facilitating nucleophilic attack from the aspartic ester amines. Consequently, reaction in acetonitrile at 80°C afforded the final compounds **28** – **30** in good yield. This provided four final compounds comprising of 1,4-dicarboxylic acid and -diester containing motifs for later evaluation. The ^1H and ^{13}C NMR spectral analysis of compound **28** follows and can be found in **Figure 5.9** on the next page.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745



Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Examination of the ^1H NMR spectra of compound **28** (**Figure 5.9a**) revealed the previously identified structural elements of the quinazoline core, in addition to new signals belonging to the 1,4-diester functionality. To avoid repetition, we focussed only on the unexplored features of compound **28**. Beginning in the upfield region, we saw the occurrence of a doublet integrating for 2H at approximately δ 3.20 ppm which we attributed to the CH_2 of the ethyl linker of the 1,4-diester electrophile (light blue). Slightly downfield from this position were two large singlets between δ 3.63 and 3.67 ppm integrating for 3H each and signifying the presence of the methyl group of the 1,4-diester (yellow). Lastly, the proton of the trisubstituted carbon centre (purple) appeared as an apparent quartet at δ 5.42 ppm and integrated for 1H.

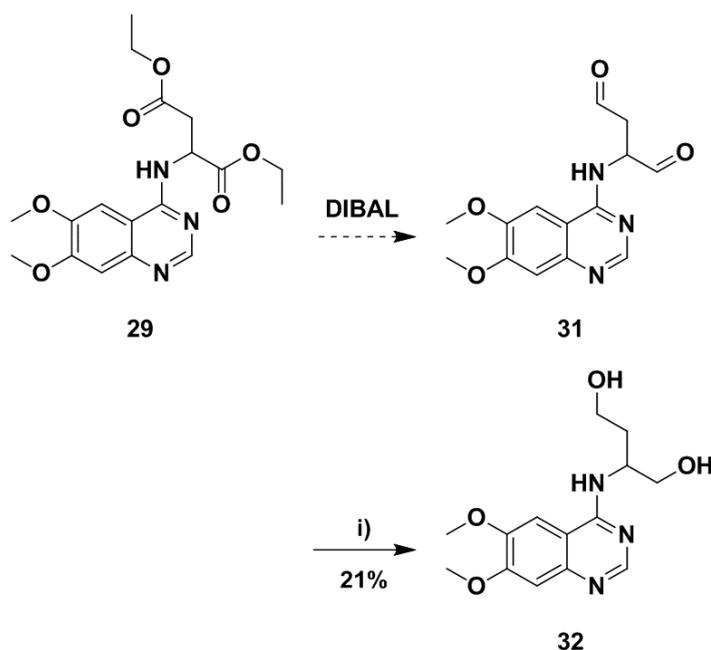
Inspection of the ^{13}C spectra in **Figure 5.9b** revealed 6 signals in the aliphatic region and 10 remaining peaks further downfield, which accounted for the 18 carbons of compound **28**. From precursor spectra and due to the electronic effects of the ester functionality, we could distinguish between the two sets of two peaks in the aliphatic region between δ 57.5 and 52.3 ppm. The peaks furthest downfield (δ 56.9 and 57.5 ppm) belonged to the quinazoline dimethoxy groups (dark blue) and the other peaks (δ 56.9 and 57.5 ppm) could be ascribed to the dimethyl ester carbons (yellow). Adjacent to these signals, we saw the peak corresponding to the trisubstituted carbon (purple) of the 1,4-diester at δ 51.3 ppm. Owing to the electron donating effect of this substitution, we expected it to be in a more downfield position than the remaining ethyl linker carbon (light blue) of the 1,4-diester, which indeed appeared at δ 35.2 ppm. In the aromatic region, we saw 8 carbon signals which correctly accounted for the quinazoline core structure. Lastly, the dicarbonyl carbons (orange) were found furthest downfield and appeared as two peaks adjacent peaks at δ 170 ppm. With all remaining protons and carbons correctly accounted for, and in conjunction with the IR and HRMS results, the discussed ^1H and ^{13}C spectra provided confirmation for the successful synthesis of compound **28**. In a similar manner, the other compounds of the series (**27**, **29** – **30**), were also validated. We therefore turned our attention towards the synthesis of the 1,4-dialdehyde electrophile-containing inhibitors.

5.4.4 Attempted 1,4-Dialdehyde Compound Synthesis

5.4.4.1 Reduction of 1,4-Diester Containing Compounds

Our first attempt towards the synthesis of the 1,4-dialdehyde-containing compounds involved reduction of the previously synthesised **29** to potentially afford compound **31**, as shown in **Scheme 5.5**. For this, we envisaged the use of diisobutylaluminium hydride (DIBAL), as an equimolar amount of the bulky reducing agent has been shown to selectively afford the aldehyde and cease further reduction at low temperature.²⁷ Unfortunately, employment of this reducing agent at controlled temperatures ranging from -78 °C to -30 °C resulted in the formation of a multitude of by-products, clearly demonstrated by TLC analysis. Consequently, purification of the reaction mixture was found to be impractical and we shifted our focus toward other synthetic routes.

One of the potential routes towards compound **31** entailed oxidation of the 1,4-diol-containing compound **32**. To obtain compound **32**, reduction of **29** was attempted using numerous reducing agents, including an excess of DIBAL, NaBH_4 , LiBH_4 and LiAlH_4 . While the greatest success was attained using LiAlH_4 , isolation and purification of the crude product proved troublesome.



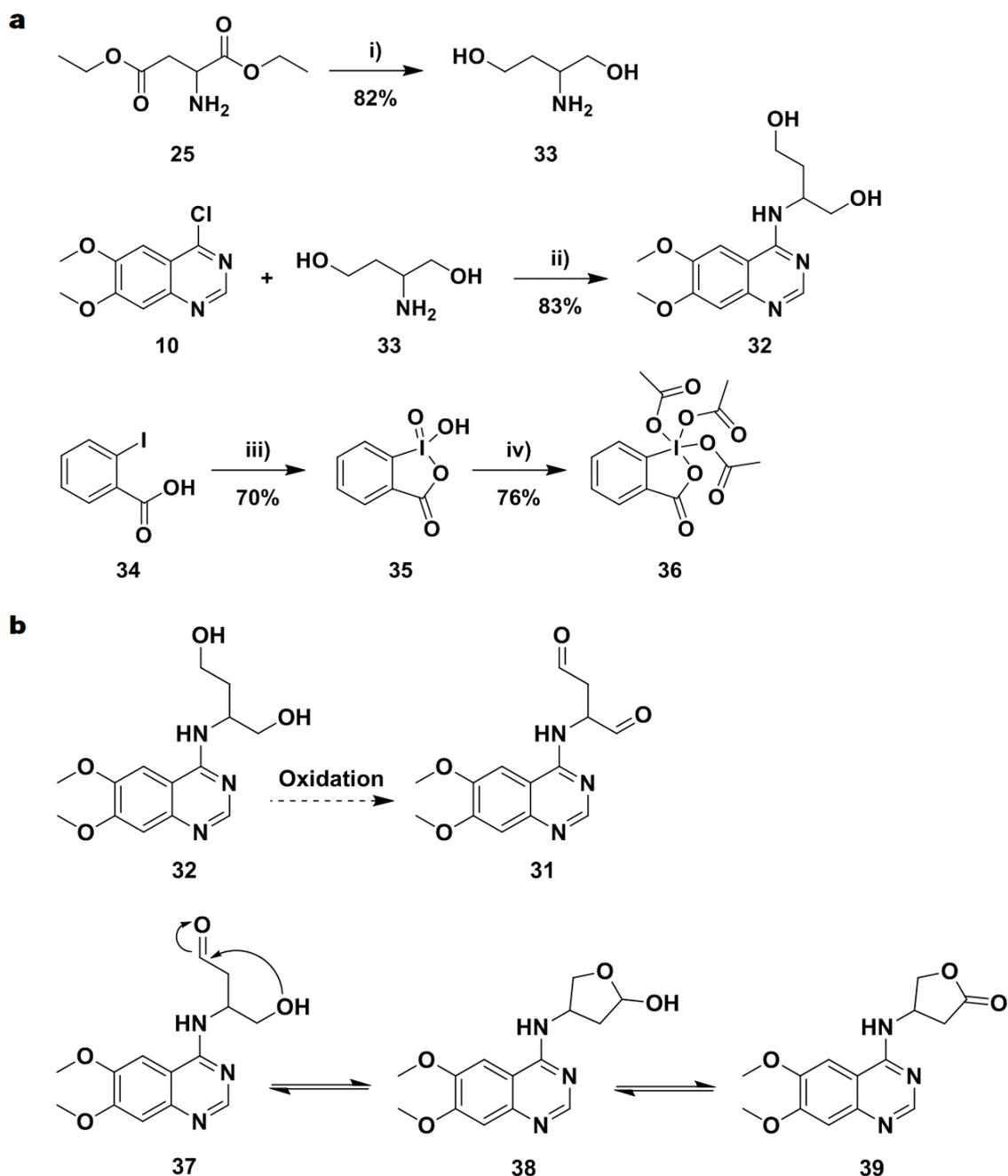
Scheme 5.5: Attempted reduction of compound **29** to 1,4-dialdehyde **31** and successful reduction to 1,4-diol **32**. Reagents and conditions: i) LiAlH_4 (3 equiv.), THF, $-10\text{ }^\circ\text{C}$ – reflux, 1 h.

A variety of work-up procedures, which included use of ethanolamine and Rochelle salt (Fieser work-up), were found to be ineffective in removing the crude from the aqueous phase.²⁸ However, this problem was circumvented by Soxhlet extraction of the solid salts, which were collected by filtration after quenching with isopropanol and water. Placement of the solid material in a Soxhlet cup, followed by continuous extraction in isopropanol while heating under reflux overnight, allowed for successful isolation of the crude product. Following the challenging purification using column chromatography, compound **32** was afforded in a low yield of 21%. While delighted with the effective isolation of this compound, we found the yield to be unsatisfactory and therefore sought a more feasible reaction sequence for the synthesis of **32** en route towards compound **31**.

5.4.4.2 Oxidation of 1,4-Diol-containing Compounds

Wishing to emulate the success of the coupling reactions of aspartic esters **24** – **26** with driving group **10** (**Scheme 5.4**), we intended to synthesise and couple compound **33** in a similar manner, as shown in **Scheme 5.6a**. Accordingly, reduction of compound **25** to **33** was carried out using LiAlH_4 with the previously optimized work up conditions followed by Soxhlet extraction. As **33** is a low molecular weight oil with a boiling point of $269\text{ }^\circ\text{C}$, purification was accomplished by Kugelrohr distillation, proving more effective than column chromatography.²⁹ Fractions with boiling point between $124\text{ }^\circ\text{C}$ – $126\text{ }^\circ\text{C}$ at approximately 20 mbar were collected, affording compound **33** in good yield. Successful coupling of this fragment to the heterocyclic driving group **10** furnished the 1,4-diol scaffold **32** in a substantially improved and workable overall yield. Preparing for the impending oxidation reactions of compound **32**, we synthesised 2-iodoxybenzoic acid (**35**) and its improved derivative Dess-Martin periodinane (**36**) from the commercially available 2-iodobenzoic acid (**34**). Both reagents have been well-documented in the oxidation of primary alcohols to aldehydes.³⁰ These were prepared according to their literature procedures and allowed for the attempted oxidation of compound **32** to the 1,4-dialdehyde **31**, as illustrated in **Scheme 5.6b**.^{31, 32}

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745



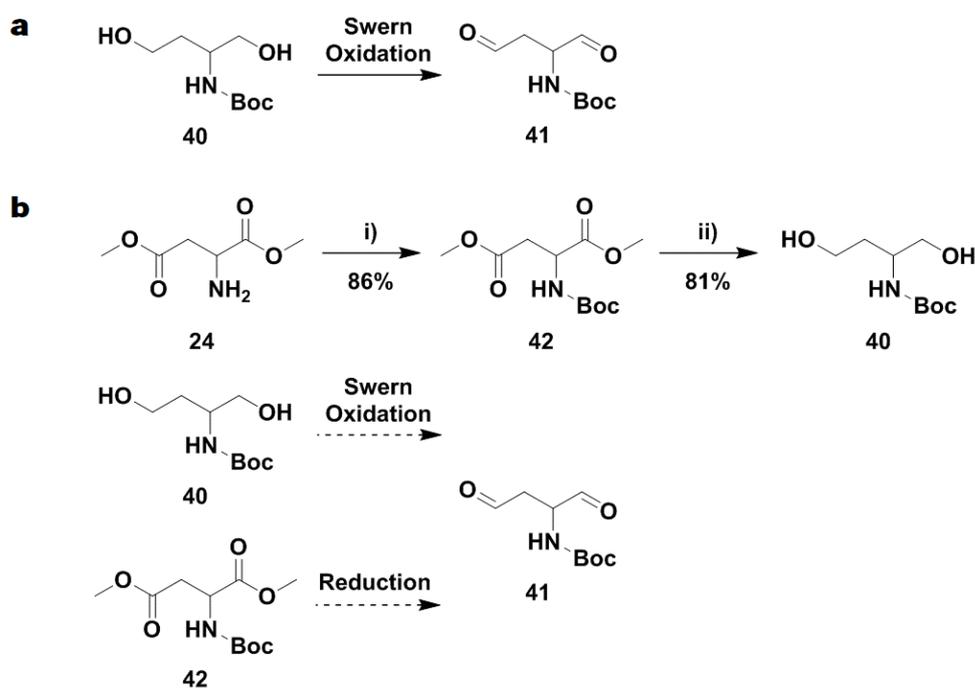
Scheme 5.6: a) Alternative synthesis of compound **32** and of oxidising agents **35** and **36**. Reagents and conditions: i) LiAlH_4 (3 equiv.), THF, $-10\text{ }^\circ\text{C}$ – reflux, 1 h; ii) 4 M HCl in dioxane (cat.), MeCN, $80\text{ }^\circ\text{C}$, 24 h; iii) Oxone (3 equiv.), H_2O , $70\text{ }^\circ\text{C}$, 3 h; iv) $\text{TsOH}\cdot\text{H}_2\text{O}$ (cat.), Ac_2O , $80\text{ }^\circ\text{C}$, 2 h. b) Attempted oxidation of compound **32** to **31** and proposed mechanism of attack by intermediates resulting in the formation of lactol **38** and lactone **39** as by-products.^{29, 31, 32}

Regrettably, the oxidation of compound **32** to **31**, and synthesis of the 1,4-dialdehyde in this manner, proved unachievable. Oxidations were attempted at varying temperatures using an array of reagents, including the synthesised **35** and **36**, pyridinium chlorochromate (PCC) and the DMSO based Pfitzner-Moffatt, Parikh-Doering and Swern oxidations with DCC, sulphur trioxide-pyridine complex and oxalyl chloride as activating species respectively.^{33, 34} However, use of these parameters invariably led to the formation of an abundance of by-products and consequently inseparable crude reaction mixtures.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

We postulated that singly oxidised intermediates such as **37** could potentially undergo nucleophilic attack from the neighbouring alcohol, resulting in formation of one of many compounds such as lactol **38** and with further oxidation lactone **39**. We were able to find two separate studies in 1,4-diol oxidation that substantiated our hypothesis. In the first of these studies, use of several oxidizing agents such as potassium permanganate, Collins reagent or PCC resulted in the formation of the γ -lactone, thought to proceed through an intermediate lactol which was readily oxidised to the corresponding lactone.³⁵ In the second study, Corey et al. were able to selectively oxidise 1,4-bis-primary alcohols to the equivalent γ -lactols using IBX.³⁶ Lastly, the quinazoline N3 and 4-amino group could potentially participate in a ring-closure reaction, similar to that seen in the formation of the imidazoquinazoline **22** in **Scheme 5.3**. For these reasons we chose to abandon this method and focus on an alternate synthetic route.

Combing through the literature, we uncovered a procedure that allowed for the oxidation of the Boc-protected compound **40** to **41** under Swern conditions, as shown in **Scheme 5.7a**.³⁷ Although the group did not isolate compound **41**, as it was immediately used in the succeeding Wittig reaction, we were encouraged by these findings. We believed that the absence of the quinazoline and presence of the Boc-protecting group could potentially prevent nucleophilic attack on the aldehyde intermediates. Following this idea, the use of di-*tert*-butyl dicarbonate with Et₃N as base in methanol furnished the intermediate **42** in good yield, as illustrated in **Scheme 5.7b**. This was followed by subsequent reduction using LiAlH₄ to afford the Boc-protected 1,4-diol **40**. Use of the documented oxidation procedure resulted in what we believed to be formation of the 1,4-dialdehyde product **41**, but to our dismay proved too unstable to isolate and verify by spectroscopic analysis. Finally, the attempted reduction of compound **42**, using DIBAL at low temperatures, succumbed to the now expected formation of by-products and inseparable crude reaction mixtures.

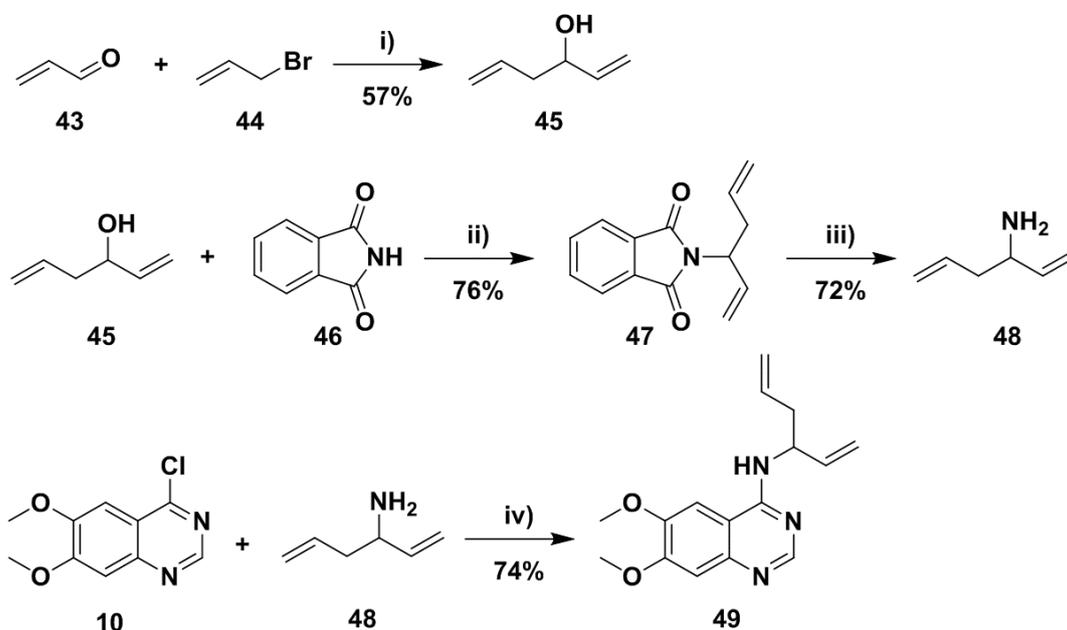


Scheme 5.7: Synthesis of Boc-protected analogues **42** and **40** and their attempted oxidation and reduction to afford compound **41**. Reagents and conditions: i) Boc₂O (1.1 equiv.), Et₃N (2.2 equiv.), MeOH, 0 °C – rt, 16 h; ii) LiAlH₄ (3 equiv.), THF, –10 °C – reflux, 1 h.^{37, 38}

5.4.4.3 Ozonolysis and Wacker Oxidation of 1,4-Diene Containing Compounds

In our final effort to produce the 1,4-dialdehyde final compound, we envisioned the synthesis of the 1,4-diene compound **49** illustrated in **Scheme 5.8**. We believed we could access the 1,4-dialdehyde or 1,4-diketone from **49** through oxidative cleavage under ozonolysis or implementation of a Wacker oxidation respectively. Hence, the synthesis of **49** commenced with a Grignard reaction between the commercially available acrolein (**43**) and allyl bromide (**44**), using magnesium turnings to generate the allylmagnesium bromide Grignard reagent *in situ*. Heating under reflux in diethyl ether, followed by purification via Kugelrohr distillation, furnished compound **45** in an acceptable yield of 57%.

Conversion of the hydroxyl group of **45** into the corresponding amine **48** was accomplished via a traditional Gabriel synthesis. Accordingly, the synthesised hydroxyl **45** and phthalimide **46** underwent a Mitsunobu reaction, employing the standard reagents of di-*tert*-butyl azodicarboxylate (DBAD) and PPh₃, affording the 1,4-diene phthalimide intermediate **47**. Subsequent aminolysis of intermediate **47** proved challenging, mainly due to the problematic isolation and purification of the product **48**. However, use of excess hydrazine hydrate in ethanol heated under reflux overnight, followed by acidic workup and filtration, allowed for successful cleavage and separation of the phthalhydrazide by-products. This afforded compound **48** as the hydrochloride salt in a yield of 72% and allowed for coupling with driving group **10** using the established conditions of catalytic acid in acetonitrile heated under reflux. The coupling proceeded smoothly, furnishing the 1,4-dialkene containing compound **49** in a good yield of 74%. The ¹H NMR spectral analysis of compound **49** follows on the next page.



Scheme 5.8: Synthesis of 1,4-diene **48** via Grignard and Mitsunobu reaction intermediates **44** and **47**. Reagents and conditions: i) **43** (1.2 equiv.), Mg turnings (1.4 equiv.), Et₂O, 45 °C, 3 h; ii) **45** (1.1 equiv.), DBAD (1.3 equiv.), PPh₃ (1.1 equiv.), THF, 0 °C – rt, 16 h; iii) hydrazine hydrate (1.5 equiv.), EtOH, 80 °C, 6 h; iv) 4 M HCl in dioxane (cat.), MeCN, 80 °C, 24 h.³⁹

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

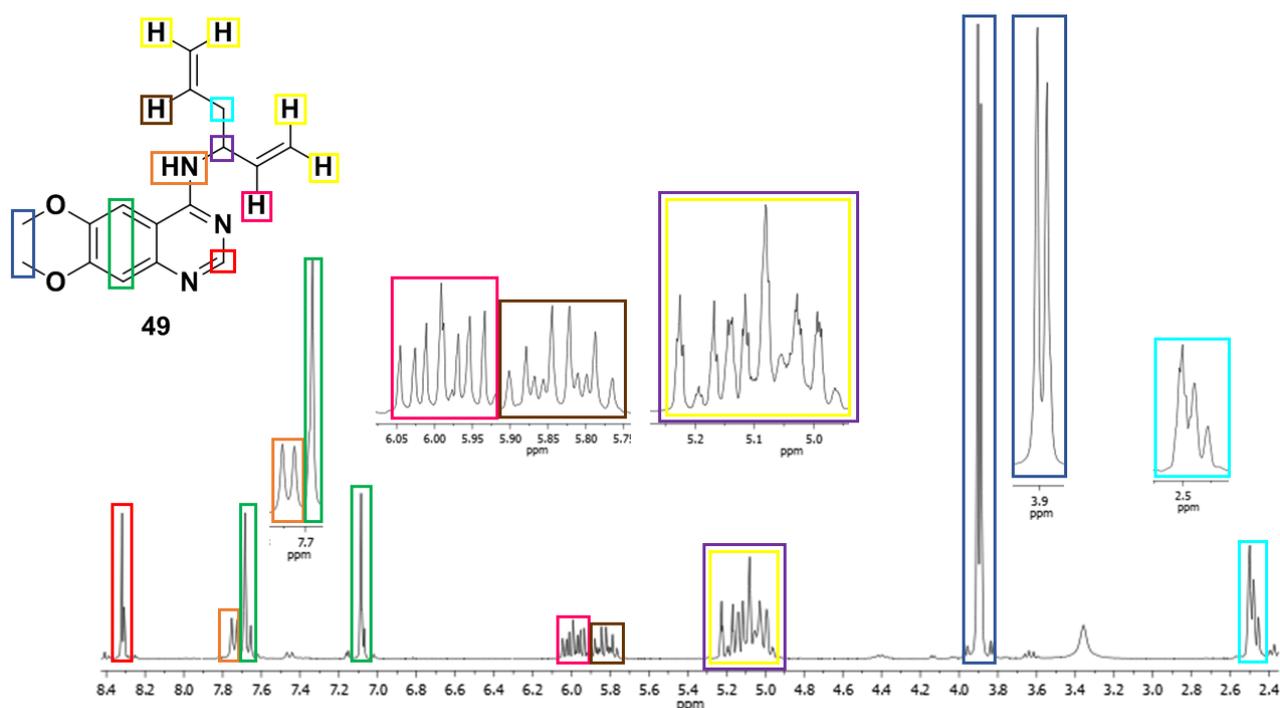
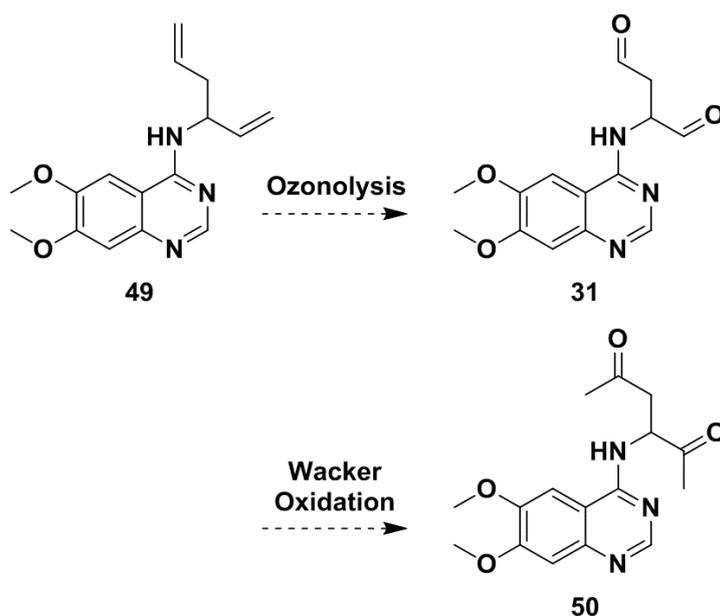


Figure 5.10: ^1H NMR spectra of final compound **49**.

Inspection of the ^1H NMR spectra of compound **49**, displayed above in **Figure 5.10**, revealed the presence of both familiar structural elements and the appearance of unique signals belonging to the 1,4-dialkene functionality. Regarding the former, we saw the signal of the ethyl linker CH_2 at δ 2.54 ppm (light blue), which appeared as a multiplet due to overlap with the $\text{DMSO-}d_6$ solvent peak. Slightly downfield from this position, we saw the aromatic methoxy groups appear as two tall singlets between δ 3.93 and 3.95 ppm, which integrated correctly for 6H. In the aromatic region, we saw 3 singlets which corresponded to the three aromatic protons of the quinazoline ring (green and red). Lastly, we found the signal of the amine proton appearing as doublet at δ 7.78 ppm (orange), in agreement with spectra of previously synthesised intermediates.

For the dialkene portion of compound **49**, we looked to the remaining signals between $\sim \delta$ 5.00 and 6.00 ppm. The splitting pattern of the set of peaks between δ 5.28 and 5.02 ppm could not be resolved and was therefore designated as a multiplet. As this multiplet integrated for 5H, we ascribed this signal to be a combination of both the terminal alkene protons (yellow) and the proton of the trisubstituted carbon centre (purple). This left the two remaining alkene protons, which appeared as two sets of signals slightly downfield. Interestingly, the nearest set at δ 5.88 ppm (brown) appeared as a doublet of doublet of triplets, with J coupling values of 17.2, 10.3 and 6.8 Hz, whilst the adjacent set of peaks at δ 6.03 ppm was resolved as a doublet of doublet of doublets, with $J = 17.2, 10.3$ and 5.8 Hz. From the predicted and observed splitting patterns it was clear that the proton adjacent to the CH_2 of the ethylene bridge should be assigned to the ddt (brown) and the remaining proton (pink) to the ddd. While not analysed, the ^{13}C spectra was correctly accounted for and may be accessed in the supplementary information (**Section 5.8**). Furthermore, HRMS analysis provided an $[\text{M}+\text{H}]^+$ ion of 286.1559, matching the calculated mass of 286.1556, which offered final confirmation for the successful synthesis of compound **49**. Therefore, with compound **49** in hand, we ventured into the ozonolysis and Wacker oxidation reactions which are shown in **Scheme 5.9**.



Scheme 5.9: Attempted ozonolysis and Wacker oxidation of compound **49** to afford **31** and **50**.

To our disappointment, the attempted ozonolysis of the 1,4-diene **49** was unsuccessful. Bubbling ozone through a solution of **49** for approximately 15 minutes at $-78\text{ }^{\circ}\text{C}$ induced the characteristic blue colour change which stems from unreacted ozone. This alluded to complete consumption of the 1,4-diene and formation of the corresponding ozonides. To convert the intermediate ozonide into the aldehyde derivative **31**, a variety of reductive work-up conditions were employed, which included use of zinc and acetic acid, dimethyl sulphide or triphenylphosphine. However, it was during this process that the reaction erupted into a slew of by-products, as visualised by TLC analysis. This once again this led to inseparable crude reaction mixtures which we believe occurred in a similar manner to that described in **Scheme 5.6b**. Lastly, we attempted the reaction in methanol with the hope of trapping the 1,4-dialdehyde as the methoxy acetal but to no avail.

Comparable outcomes were found in the attempted Wacker oxidation of compound **49** to obtain the 1,4-diketone **50**. Use of the Tsuji-Wacker protocol, with PdCl_2 as catalyst and a stoichiometric amount of CuCl in a mixture of DMF and water, resulted in consumption of the starting material and formation of numerous products.⁴⁰ This made the subsequent separation and purification near impossible. Furthermore, utilisation of $\text{Pd}(\text{OAc})_2$ and trifluoroacetic acid was found to be similarly ineffective.⁴¹

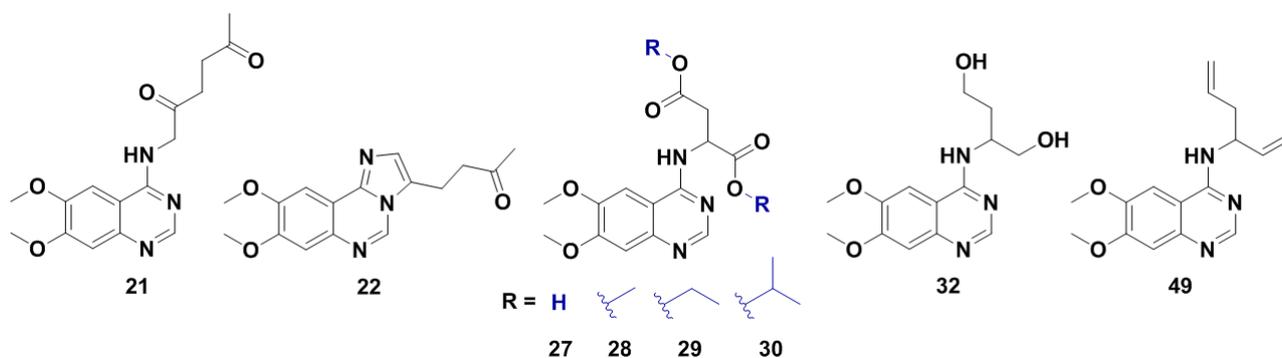
Several alternative synthetic strategies were employed in the attempted synthesis of 1,4-dialdehyde final compounds and their acetal protected equivalents. While these routes led to the synthesis of assorted intermediate compounds, none proved successful and these attempts will therefore not be elaborated on. At this point it was decided to evaluate the inhibitory properties of the successfully synthesised final compounds. From the relevant biochemical and cellular activity, we wished to ascertain whether it would be worthwhile proceeding in the synthesis of this class of compounds. We elected to include the 1,4-diol and 1,4-diene-containing compounds **32** and **49** respectively to expand the scope of the compounds structure activity relationship. This meant that a total of 8 final compounds were evaluated against the wild type and mutant variant forms of EGFR and follows in the next section.

5.5 Biochemical and Cellular Evaluation

To determine efficacy in a biochemical setting, the synthesised compounds were evaluated against the wild-type, activating L858R and double L858R/T790M mutant variants of the enzyme. For the cellular tests, we employed a wild type-bearing cell line A431 and the EGFR-L858R/T790M drug resistant cell line H1975. Compound assessment was carried out by Dr. Jonas Lategahn, with all measurements being executed in at least triplicate. As the synthesised compounds are quinazoline based, the respective IC₅₀ and EC₅₀ values of gefitinib accompanied the obtained results.

An overview of the results from the biochemical and cellular evaluation of the compounds synthesised in this chapter can be seen below in **Table 5.1**. Inspection of the cellular results (CTG) revealed all compounds showing no activity below the limit of 30 μ M. This could be ascribed to several oversights in the fundamental compound design, the first of which is the lack of a solubilising group. We underestimated the significance of including the solubilising group, which is a conventional attribute in kinase inhibitor drug design.

Table 5.1: Biochemical and cellular evaluation results for the synthesised 1,4-dicarbonyl final compounds.



Compound	EGFR HTRF IC ₅₀ [nM]			EGFR CTG EC ₅₀ [μ M]	
	wild-type	L858R	L858R/T790M	A431	H1975
Gefitinib	0.37 \pm 0.76	0.15 \pm 0.25	189 \pm 103	1.71 \pm 0.79	10.7 \pm 2.55
21	10000 \pm 0	10000 \pm 0	10000 \pm 0	30 \pm 0	30 \pm 0
22	7030 \pm 4198	870 \pm 259	10000 \pm 0	30 \pm 0	30 \pm 0
27	10000 \pm 0	10000 \pm 0	10000 \pm 0	30 \pm 0	30 \pm 0
28	587 \pm 86.3	61.26 \pm 15.1	6652 \pm 4345	30 \pm 0	30 \pm 0
29	2351 \pm 365	185 \pm 30.8	10000 \pm 0	30 \pm 0	30 \pm 0
30	8482 \pm 2629	515 \pm 147	10000 \pm 0	30 \pm 0	30 \pm 0
32	1300 \pm 456	17.3 \pm 9.33	10000 \pm 0	30 \pm 0	30 \pm 0
49	17.3 \pm 8.01	1.97 \pm 2.79	356 \pm 137	30 \pm 0	30 \pm 0

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Its inclusion plays an important role not only in the pharmacokinetic profile and overall solubility of the compound within the cell, but also in anchoring the inhibitor within the active site. This allows for a favourable orientation of the molecule and thus improved reversible interactions with the hinge region. Furthermore, these results reflect the findings of the previous chapter, when comparing the compounds synthesised with and without a solubilising group.

In hindsight, we believe that several of the included 1,4-dicarbonyl warheads negatively affected the overall pharmacokinetic dynamics of the synthesised inhibitors. As an example, compound **27** contains a dicarboxylic acid moiety, a functional group that has been shown to promote metabolic instability, toxicity and limited passive diffusion across biological membranes in medicinal applications.⁴² Moreover, the other 1,4-dicarbonyl compounds such as the 1,4-diester compounds **28** – **30** may hydrolyse within the cell to the carboxylic acid equivalents.

While the above discussed factors do translate negatively to the biochemical activity (**HTRF**) of the compounds, mainly by influencing reversible binding affinity, we do see some noteworthy IC_{50} activities. Surprisingly, we see inhibition over the limiting threshold of 10 μM for the 1,4-diketone compound **21** which is therefore considered inactive. We believe the root of this inactivity to lie in a steric clash between the 1,4-diketone chain and the $\beta 3$ sheet or helix αC strand at the roof of the enzyme. This is suggested from the length and linear character of the 1,4-diketone chain, as well as the angle from which it extends at the 4-position of the quinazoline scaffold. Our posit is further validated by a comparison with the ring closed counterpart compound **22**, which shows limited activity against the wild-type and sub-micromolar IC_{50} values against the single activating EGFR-L858R mutant. Counting from the quinazoline ring, compound **22** is two carbons shorter than **21** with a far more horizontal angle of extension, which could avoid a potential steric clash.

Compound **22** displays a desirable selectivity profile, with an 8-fold increase in activity against the L858R mutant variant over the wild-type. Visual inspection of the activity table reveals this selectivity to be a recurrent trend amongst all compounds that displayed biochemical inhibitory activity. Furthermore, all compounds except **28** and **49** were considered inactive against the L858R/T790M double mutant. This however was to be expected, as quinazoline-based 1st generation inhibitors such as gefitinib and erlotinib were specifically identified and designed to target the activated EGFR-L858R enzyme. Additionally, and further validating this anticipated outcome, the L858R/T790M double mutant was responsible for bringing about resistance to 1st generation EGFR inhibitors through the steric bulk of the methionine residue and increased affinity to ATP. Both of these concepts were discussed in **Section 2.4.1** and **Section 2.4.2** respectively. As this study was focussed on the possibility of covalently modifying the catalytic lysine residue within the active site, we were unperturbed by the lack of activity against the L858R/T790M double mutant.

Compound **27** was shown to be completely inactive across the board. As previously discussed, we believe this to be due to a poor pharmacokinetic profile, with the double carboxylic acid moieties disrupting the ability for the inhibitor to bind reversibly. The 1,4-diester compounds **28** – **30** all showed activity against the wild-type and L858R mutant. Closer examination of these activities displays a clear trend, with a decrease in activity accompanying a more bulkier ester substituent. In near linear fashion, the ethyl (**29**) and isopropyl (**30**) ester were outperformed by the methyl ester **28**, with an IC_{50} value of 61 nM against EGFR-L858R. This alludes to the limited space in the hydrophobic pocket and potential steric clashes in this region, as mentioned for compound **21**. Furthermore, as with compound **21** and **22**, the horizontal, branched 1,4-dicarbonyl chain systems proved superior to longer and linear chains.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

To our astonishment, the final two compounds **32** and **49** proved to be the most efficacious of the inhibitors evaluated. The 1,4-diol containing compound **32** exhibited an IC_{50} value of 17 nM against the L858R activating mutant, with an outstanding 75-fold selectivity over the wild-type. We postulated that the enhanced activity of this compound arises from several factors. Firstly, the absence of a bulky carbonyl substituent at the terminal end of the branched chain eliminated the possibility of a steric clash within the active site. Additionally, it appears the hydrogen bond donor features of these terminal hydroxyl moieties are more favourable than the hydrogen bond acceptor character of the 1,4-dicarbonyl counterparts. This could be due to the formation of reversible interactions with either the Glu-Lys salt bridge or the nearby glycine rich loop. These proposed interactions prompted us to speculate on the compound's selective inhibition of the activating mutant. The L858R mutation induces a constitutively activated form of the receptor, leading to the sustained formation of the Glu-Lys salt bridge and a DFG-in motif as explained in **Section 5.2.1**. In these circumstances, the 1,4-diol compound **32** would be able to form the aforementioned reversible interactions, whilst with the inactive wild-type enzyme it would not. This provides a potential explanation for the exceptional selectivity profile.

The 1,4-diene compound **49** displayed a single digit activity of approximately 2 nM against the L858R-EGFR variant, with a nearly 9-fold selectivity over the wild-type – similar to that found for compounds **22** and **28 – 30**. Interestingly, **49** also demonstrated good activity against the L858R/T790M double mutant, with an IC_{50} value of 356 nM. Activity against this variant was unobserved in any of the other synthesised compounds. Owing to the difficulty in explaining the gain in activity for this compound, the biochemical and cellular evaluation for this compound were undertaken 5 times to ensure the correct IC_{50} values. As with **32**, we believe the smaller terminal alkene functionality of compound **49** would likely aid in eliminating potential steric clashes within the active site. This feature, coupled with the binding mode of the compound, could be particularly favourable in avoiding the bulky methionine residue of EGFR-T790M. Furthermore, the results indicate that the non-polar, electron-rich olefins could be forming π -interactions with specific residues. With the available information in hand this is unsubstantiated and any further attempts to clarify this enhanced activity would be conjecture. However, this remains an intriguing result and it is our endeavour to investigate this functional group and the source of this activity further. This examination is outlined within the future work section (**Section 5.7**).

To conclude, confirmation of the compounds binding properties and mode of inhibition was investigated using mass spectrometry experiments. For this, wild-type EGFR was incubated with the synthesised final compounds and DMSO was used as control. Unfortunately, none of the incubated enzymes resulted in a mass increase equivalent to the corresponding compound labelled protein and the masses obtained were found to be comparable to the control wild-type EGFR treated with DMSO. These results indicated a reversible binding mode for all synthesised compounds, corroborating with the biochemical and cellular data and thereby ruling out covalent bond formation with the targeted catalytic lysine.

5.6 Conclusions

The research outputs described in this chapter focussed on exploring the viability of covalent bond formation between a 1,4-dicarbonyl electrophile and the EGFR catalytic lysine residue. To assess this feasibility, we successfully synthesised a gefitinib-based heterocyclic driving group derivatised with 1,4-diketone, 1,4-dicarboxylic and 1,4-diester electrophiles to potentially undergo covalent modification. Despite numerous attempts and synthetic strategies, the 1,4-dialdehyde counterpart could regrettably not be produced. We believe this stems from reactivity complications inherent to the dialdehyde system, which is particularly relevant to the linear and branched alkyl systems we attempted to synthesise. The sp^3 hybridisation of these compounds allowed for rotational freedom and proximity of the reactant groups, which in turn expedited intramolecular reactions and formation of numerous by-products. To overcome this, we envisage the use of either aromatic or cyclic derived 1,4-dialdehydes, as these are the only systems which have found success in the literature, as discussed in **Section 5.2.2**. Prospective studies in the use of these electrophiles and their synthesis will be discussed in detail in the following section.

Evaluation of the synthesised compounds revealed oversights taken on our part and some unexpected outcomes, the first being inactivity of all compounds in a cellular context. We strongly believe this to be due to a poor pharmacokinetic profile, elicited mainly by the lack of a solubilising group, as the compounds retained some activity biochemically. Therefore, we will be including a solubilising group within the structure in future studies concerning this class of inhibitor. A comprehensive procedure for its inclusion can be found in the next section.

We were able to identify several trends from the biochemical data. Generally, a reduction in steric bulk led to greater inhibitory activity, with shorter, non-linear chains proving the most effective. This was evident in the 1,4-diketone **21** and the three 1,4-diester compounds **28** – **30**. This tendency extended to the 1,4-diol and 1,4-diene compounds **32** and **49** respectively, produced en route to the 1,4-dialdehyde, where the smaller terminal functional groups were able to evade steric clashes. Compound **49** showed an atypically superior activity profile against both the single and double mutant variants of EGFR and demands further investigation to understand the underlying mechanisms of its efficacy. To retain and transfer the identified desirable SAR characteristics, we envisage incorporation of an aromatic group at the 4-position of the driving group, similar to that found in gefitinib, with the appropriate electrophile attached. Elaboration and exposition of this, and how compound **49** may be explored further, is conferred in the following future work section (**Section 5.7**).

Covalent mass spectrometry studies disclosed that all the synthesised inhibitors bound to the enzyme in a reversible manner. Failure to promote covalent modification can be ascribed to two factors. Firstly, we maintain that the synthesised electrophiles were unable to achieve an optimal position and orientation to undergo covalent bond formation with the catalytic lysine residue. This is attributed to the previously discussed electrophile chain length/structure and steric interactions within the active site. Secondly, we suspect that the electrophilic reactivity of the synthesised 1,4-dicarbonyl fragments was inadequate to allow for irreversible inhibition. From its inception, this research endeavour has focussed on the synthesis of the 1,4-dialdehyde-containing compounds as the primary objective, with the other 1,4-dicarbonyl electrophiles as a supplementary for reactivity comparison.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Thus, we believe that the synthesis of this functionality and its subsequent screening will provide the best platform to substantiate this claim and the feasibility of this study. This is addressed as a large portion of the future work (**Section 5.7**).

Targeting lysine residues with irreversible inhibitors is an emerging but burgeoning field of medicinal chemistry, where a high degree of serendipity is often required to achieve covalent bond formation.¹² With increasing examples of its use, and improved knowledge of optimal lysine-selective electrophile development and required reactivity, the factors determining successful lysine-targeting covalent inhibition will become more understood and appreciated. This will lead to broader application and utilisation within the drug discovery process.

While we maintain that the strategy of lysine-targeting covalent inhibition described in this chapter holds potential, we elected to postpone further investigation of this approach. Concurrent to the completion of this body of work, the first reported crystal structure of EGFR, in which the catalytic lysine is covalently modified, was published.⁴³ It is upon the premise of this publication that the succeeding chapters research is based. This leads directly to the following chapter, where this publication and our new research endeavour will be comprehensively discussed.

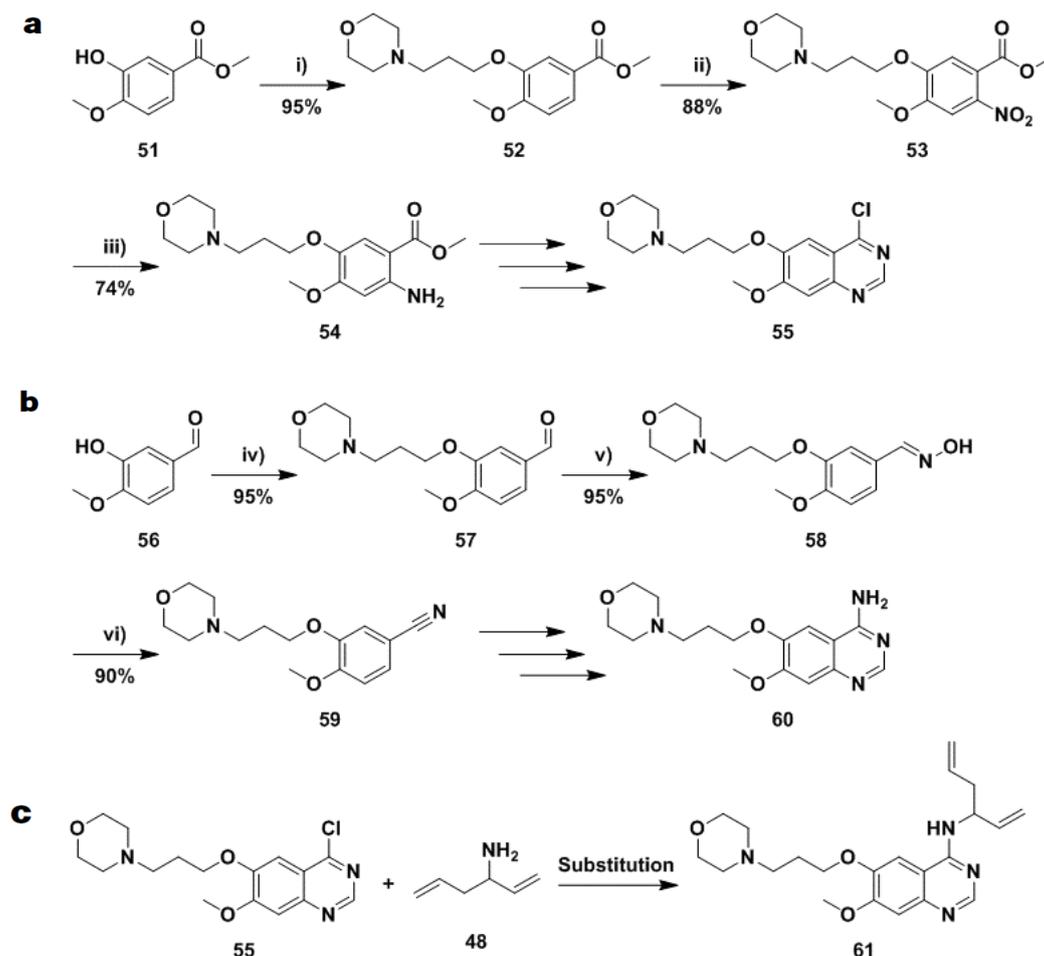
5.7 Future Work

Owing to the synthesised 1,4-dicarbonyl compounds poor performance in their cellular assessment, the first step in improving these values will be through the inclusion of a solubilising group. We believe this will drastically improve the pharmacokinetic profile and binding affinity by anchoring the scaffold and allowing for stronger reversible interactions. In **Scheme 5.10** on the following page, we have outlined a synthesis for the introduction of a propyl linked morpholine solubilising group, the same that is utilised in gefitinib. All yields shown in schemes within this section are those provided within the relevant literature procedures.

A coupling reaction at the 3-hydroxyl of the commercially available **51** with 4-(3-chloropropyl)morpholine and K_2CO_3 serving as base will provide compound **52**, as shown in **Scheme 5.10a**. This will be followed by selective nitration (**53**) and reduction to form the anthranilic ester derivative **54**. Subsequent reactions can then be performed in a similar manner to that described in **Scheme 5.1**, affording the 4-chloro heterocyclic scaffold **55** with a solubilising group. Looking at **Scheme 5.10b**, a similar coupling will be carried out on isovanillin to produce compound **57**, after which the nitrile-containing compound **59** may be synthesised via formation of oxime **58** using hydroxylamine hydrochloride. Once again, subsequent reactions to afford the heterocyclic driving group **60** will be performed in a similar manner to that described in **Scheme 5.1**.

Following the successful synthesis of driving groups **55** and **60**, pursuit of compound **61**, illustrated in **Scheme 5.10c**, could provide insight into the unexplained potency of our most efficacious inhibitor **49**. This compound can be easily furnished through coupling of **48** and **55** under the previously established conditions. Addition of the solubilising group could not only impart a potential increase in inhibitory efficacy, but also allow for incubation and co-crystallisation with the L858R single mutant enzyme due to the improved binding affinity. This would paint an unequivocal picture and explanation pertaining to the compound's unforeseen effectiveness.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

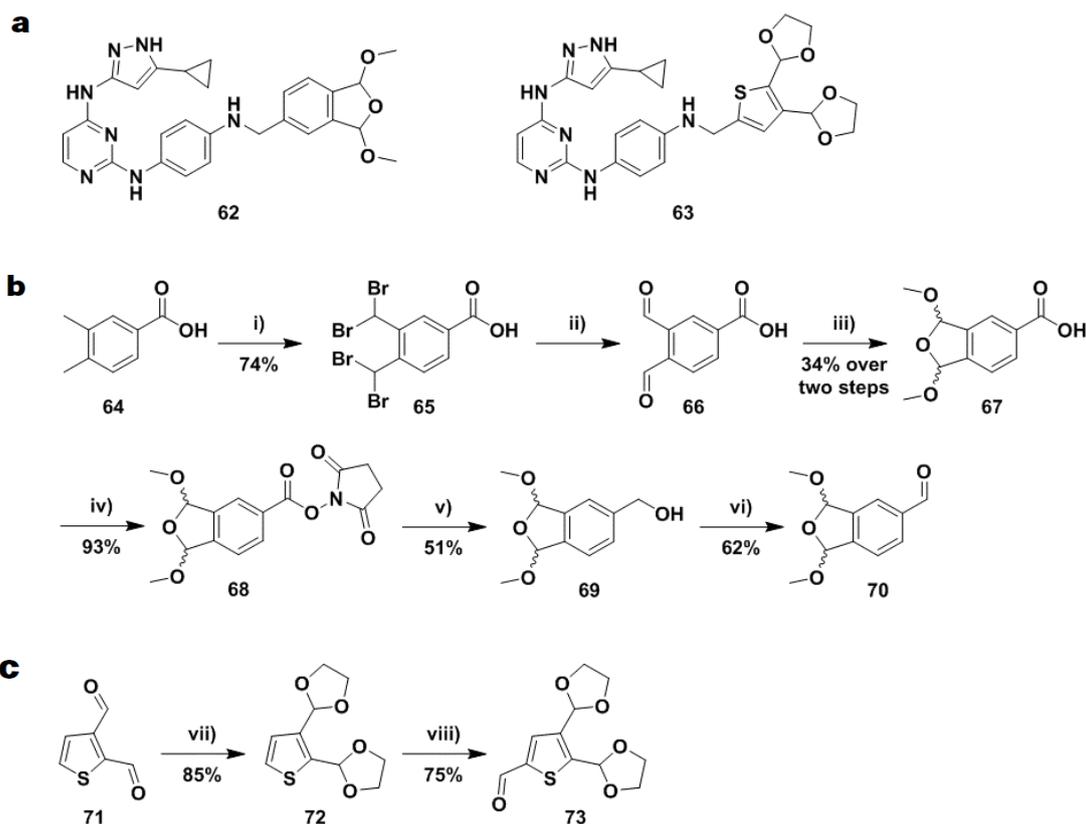


Scheme 5.10: a) Proposed synthesis of the heterocyclic driving group scaffold **55** and b) **60** with gefitinib solubilising group and c) Proposed synthesis of compound **61** for further biochemical and cellular investigation. Reagents and conditions: i) 4-(3-chloropropyl)morpholine, K_2CO_3 , TBAI, 2-butanone, 80 °C; ii) HNO_3 , Ac_2O , -10 °C – rt; iii) H_2 , Pd/C, EtOH:EtOAc; iv) 4-(3-chloropropyl)morpholine, K_2CO_3 , DMF, 100 °C; v) $NH_2OH \cdot HCl$, pyridine, MeOH, 70 °C; vi) Ac_2O , 100 °C.^{44, 45}

In a similar approach to that described in **Section 5.2.2** and illustrated in **Figure 5.5a**, Statsuk et al. developed a three-component chemical reaction for trapping kinase substrate complexes.⁴⁶ Initial failures using a weakly reversible binding adenosine scaffold, which resulted in nonspecific side reactions, led to replacement with the potent kinase inhibitor scaffold **62**, shown in **Scheme 5.11a**. This compound contains an acetal protected *o*-phthalaldehyde fragment which served as the lysine trap. Furthermore, the group were also able to use the less reactive acetal protected thiophene-2,3-dicarboxaldehyde containing compound **63**, to effectively cross-link with cysteine-containing substrates.

Based on this work, we wish to incorporate the *o*-phthalaldehyde and thiophene-2,3-dicarboxaldehyde moiety into our heterocyclic driving group at the 4-position. Inclusion of these functionalities can be beneficial in several ways. Stability is imparted by the attachment of these groups to aromatic systems, which assists in counteracting the unwanted formation of intramolecular by-products. Furthermore, an aromatic substitution at the 4-position mimics that which is found in gefitinib, potentially providing an optimal position and orientation for the electrophiles to covalently modify the catalytic lysine. Lastly, use of the acetal-protected counterpart allows for a more convenient synthetic route and provides easy access to the desired 1,4-dialdehyde electrophile through a final deprotection.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745



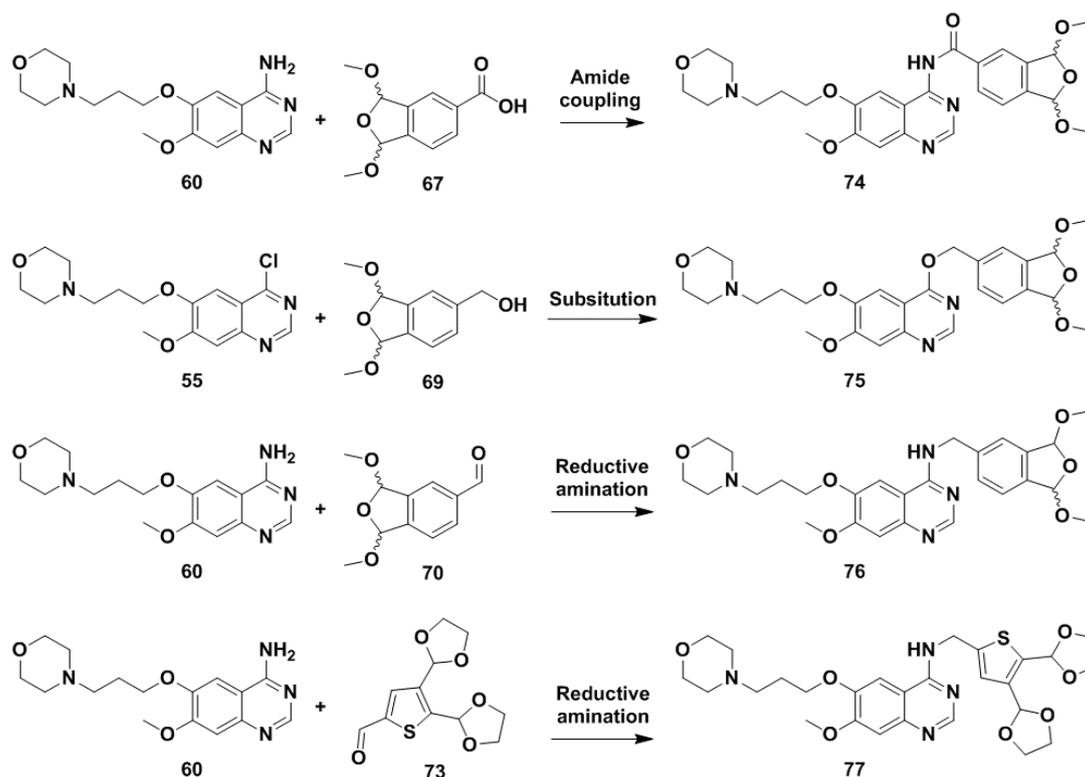
Scheme 5.11: a) Kinase substrate trapping complexes **62** and **63**, b) Proposed synthesis of the acetal protected phthalaldehyde fragments **67**, **69** and **70** and c) Proposed synthesis of the thiophene counterpart **73**. Reagents and conditions: i) NBS, CCl_4 , 80°C ; ii) Na_2CO_3 , H_2O , 70°C ; iii) $\text{Sc}(\text{OTf})_3$, MeOH, rt; iv) NHS, DCC, MeCN, rt; v) NaBH_4 , THF, rt - 70°C ; vi) COCl_2 , Et_3N , DMSO, -78°C ; vii) ethylene glycol, $\text{TsOH}\cdot\text{H}_2\text{O}$ (cat.), benzene, 80°C ; viii) BuLi, DMF, THF, -78°C – rt.⁴⁶

Synthesis of the acetal protected *o*-phthalaldehyde fragment **70** proceeds through several useful intermediates and can be seen in **Scheme 5.11b**.⁴⁶ Benzylic bromination of the commercially available **64** affords the tetra-brominated compound **65**, which may then undergo subsequent hydrolysis to the dialdehyde (**66**), followed by acetal formation using scandium triflate as the highly active Lewis acid to furnish intermediate **67**. Amide coupling with *N*-hydroxysuccinimide using *N,N'*-dicyclohexylcarbodiimide will afford compound **68**, setting up for the successive reduction to the alcohol intermediate **69**. Lastly, compound **69** may be oxidised to the 5-position aldehyde-containing *o*-phthalaldehyde **70** using a Swern oxidation.⁴⁶

The synthesis of the thiophene counterpart **73** is illustrated in **Scheme 5.11c** above.⁴⁶ Initial acetal protection of the commercially available **71** with ethylene glycol, making use of a Dean-Stark apparatus to eliminate water, will provide the acetal protected equivalent **72**. This is followed by formylation at the activated 2-position using butyl lithium and quenching with DMF, affording fragment **73** for future use.

We envisage producing the final compounds shown in **Scheme 5.12** through a variety of reactions between driving groups **55** and **60** and the synthesised fragments **67**, **69** – **70** and **73**. Amide coupling between **60** and **67**, potentially using a carbodiimide reagent and additive, could afford the 4-position amide compound **74**. Substitution between the electrophilic 4-chloro quinazoline **55** and alcohol nucleophile **69**, with K_2CO_3 or Et_3N as base, may serve to furnish the ether linked final compound **75**.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

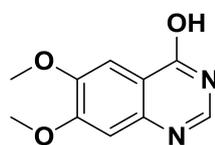


Scheme 5.12: Proposed synthesis to afford the acetal protected 1,4-dialdehydes **74** - **77**.

Lastly, reductive amination between the 4-amino quinazoline **60** and aldehydes **70** and **73**, using sodium cyano- or triacetoxy-borohydride, could furnish the desired amine-linked acetal protected *o*-phthalaldehyde and thiophene-2,3-dicarboxaldehyde compounds **76** and **77** respectively. It is important to note that all the synthesised acetal-protected compounds may also be deprotected to their 1,4-dialdehyde counterparts. This can be carried out using TFA, as achieved by Shokat and co-workers in their kinase substrate crosslinking experiments.⁴⁶ Synthesis of these compounds could provide valuable insight pertaining to the targeted covalent inhibition of the EGFR catalytic lysine residue.

5.8 Supplementary Information

6,7-Dimethoxyquinazolin-4-ol (**9**)⁴⁷

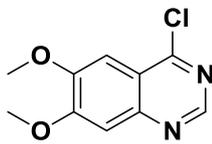


A 2-neck round-bottom-flask was charged with methyl 2-amino-4,5-dimethoxybenzoate (5.00 g, 23.7 mmol, 1 equiv.), ammonium formate (2.24 g, 35.5 mmol, 1.5 equiv.), purged with Ar and dissolved in formamide (50 mL). The solution was then allowed to stir at 140 °C for 8 h or until complete consumption of the starting material as indicated by TLC, during which time the solution changed from orange to yellow in colour. After allowing to cool to rt, the solution was then poured into ice-cold H₂O (100 mL) resulting in the formation of a light-brown precipitate. The precipitate was then collected by filtration and washed with aliquots of cold H₂O (3 × 40 mL) and dried under vacuum to afford compound **9** as an off-white solid (4.40 g, 21.3 mmol, 90%) with no further purification required.

¹H NMR and MS data collected for this compound compared well with the reported literature values.⁴⁷

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

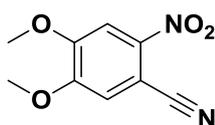
Rf: 0.26 (5% MeOH in DCM); mp: 296 – 298 °C; ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 3.86 (s, 3H, ArOCH_3), 3.90 (s, 3H, ArOCH_3), 7.12 (s, 1H, ArH), 7.43 (s, 1H, ArH), 7.98 (s, 1H, ArH), 12.07 (br s, 1H, ArOH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 55.7, 55.9, 104.9, 108.0, 115.6, 143.9, 144.9, 148.5, 154.4, 160.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{10}\text{H}_{11}\text{N}_2\text{O}_3$: 207.0770; found: 207.0769.

4-Chloro-6,7-dimethoxyquinazoline (10)⁴⁸

A flame-dried, 2-neck round-bottomed flask was charged with 6,7-dimethoxyquinazolin-4-ol (1.00 g, 4.85 mmol, 1 equiv.), purged with Ar and to this was slowly added an excess of thionyl chloride (7.00 mL, 97.0 mmol, 20 equiv.). To the resulting solution was added a catalytic amount of DMF (0.110 mL, 1.46 mmol, 30 mol%) dropwise and the reaction mixture was then allowed to stir at 80 °C for 6 h. The excess thionyl chloride was removed by vacuum distillation and the residue was then diluted with DCM (50 mL) and H_2O (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 \times 20 mL) and the combined organic layers were then washed with a saturated solution of NaHCO_3 (3 \times 20 mL), brine (20 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved by recrystallisation with DMF to afford compound **10** as white crystals (959 mg, 4.27 mmol, 88%).

^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{47, 48}

Rf: 0.66 (5% MeOH in DCM); mp: 184 – 186 °C; ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 3.98 (s, 1H, ArOCH_3), 3.98 (s, 1H, ArOCH_3), 4.00 (s, 1H, ArH), 7.34 (s, 1H, ArH), 7.40 (s, 1H, ArH), 8.85 (s, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 56.1, 56.5, 102.1, 106.7, 118.5, 148.5, 151.3, 152.1, 156.6, 157.8 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{10}\text{H}_{10}\text{N}_2\text{O}_2\text{Cl}$: 225.0431; found: 225.0427.

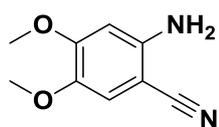
4,5-Dimethoxy-2-nitrobenzonitrile (12)²³

A flame-dried, 2-neck round-bottomed flask was charged with 3,4-dimethoxybenzonitrile (4.00 g, 24.5 mmol, 1 equiv.), purged with Ar and dissolved in acetic anhydride (35 mL). The resulting solution was cooled to –10 °C (acetone/ice) and this was followed by the dropwise addition of nitric acid (10.2 mL, 245 mmol, 10 equiv.) over 20 min. The reaction mixture was then allowed to warm to rt and stirred for 12 h or until complete consumption of the starting material as indicated by TLC. The solution was then poured into ice-cold H_2O (150 mL) resulting in the formation of a yellow precipitate. The precipitate was then collected by filtration and washed with aliquots of H_2O (3 \times 50 mL) and dried under vacuum to afford compound **12** as a yellow solid (5.05 g, 24.3 mmol, 99%) with no further purification required.

^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{23, 49, 50}

Rf: 0.46 (30% EtOAc in PE); mp: 167 – 168 °C; ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 3.96 (d, J = 1.4 Hz, 6H, 2 \times ArOCH_3), 7.68 (s, 1H, ArH), 7.86 (s, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 56.7, 57.1, 99.6, 108.4, 115.8, 116.4, 142.3, 151.8, 153.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_9\text{H}_9\text{N}_2\text{O}_4$: 209.0562; found: 209.0560.

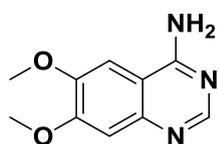
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

2-Amino-4,5-dimethoxybenzonitrile (13)²³

A 2-neck round-bottom flask was charged with 4,5-dimethoxy-2-nitrobenzonitrile (5.00 g, 24.0 mmol, 1 equiv.), purged with Ar and dissolved in chloroform (70 mL). To this was added tetrabutylammonium bromide (1.93 g, 6.00 mmol, 25 mol%) in one portion, followed by the slow addition of a solution of sodium dithionite (33.4 g, 192 mmol, 8 equiv.) in H₂O (70 mL). The reaction mixture was then allowed to stir at rt for 16 h, basified to pH = 9 with a solution of 2 M NaOH and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 50 mL) and the combined organic layers were concentrated to ~ 30 mL. Whilst stirring gently, a solution of 4 M HCl in dioxane was added slowly until precipitation occurred and the resulting precipitate was collected by filtration and dissolved in 2 M NaOH solution (50 mL). The aqueous layer was extracted with aliquots of DCM (3 × 50 mL) and the combined organic layers were then washed with a saturated solution of NaHCO₃ (30 mL), brine (30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, compound **13** was afforded as a yellow solid (2.39 g, 13.4 mmol, 56%) with no further purification required.

¹H NMR and MS data collected for this compound compared well with the reported literature values.^{23, 51}

Rf: 0.58 (60% EtOAc in PE); mp: 99 – 100 °C; ¹H NMR (400 MHz, DMSO-*d*₆) δ 3.64 (s, 3H, ArOCH₃), 3.73 (s, 3H, ArOCH₃), 5.61 (s, 2H, ArNH₂), 6.41 (s, 1H, ArH), 6.87 (s, 1H, ArH) ppm; ¹³C NMR (101 MHz, DMSO-*d*₆) δ 55.3, 56.2, 83.5, 98.9, 114.0, 118.7, 140.3, 148.4, 154.4 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₉H₁₁N₂O₂: 179.0821; found: 179.0823.

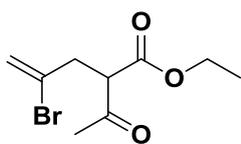
6,7-Dimethoxyquinazolin-4-amine (14)⁴⁷

A 2-neck round-bottom flask was charged with 2-amino-4,5-dimethoxybenzonitrile (2.20 g, 12.3 mmol, 1 equiv.), ammonium formate (1.17 g, 18.5 mmol, 1.5 equiv.), purged with Ar and dissolved in formamide (25 mL). The solution was then allowed to stir at 140 °C for 10 h or until complete consumption of the starting material as indicated by TLC. After allowing to cool to rt, the solution was then poured into ice-cold H₂O (50 mL) and left overnight to allow for precipitation. The precipitate was then collected by filtration and washed with aliquots of cold H₂O (3 × 40 mL) and dried under vacuum to afford compound **14** as a tan solid (2.37 g, 11.6 mmol, 94%) with no further purification required.

¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.⁵²

Rf: 0.46 (10% MeOH in DCM); mp: 206 – 208 °C; ¹H NMR (300 MHz, DMSO-*d*₆) δ 3.86 (s, 3H, ArOCH₃), 3.89 (s, 3H, ArOCH₃), 7.07 (s, 1H, ArH), 7.41 (s, 2H, ArNH₂), 7.57 (s, 1H, ArH), 8.26 (s, 1H, ArH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 55.7, 56.0, 102.6, 106.8, 108.0, 146.6, 148.2, 153.9, 154.0, 160.4 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₀H₁₂N₃O₂: 206.0930; found: 206.0930.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

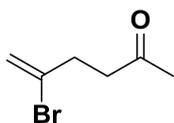
Ethyl 2-acetyl-4-bromopent-4-enoate (17)⁵³

A flame-dried, 2-neck round-bottomed flask was purged with Ar and to this was added anhydrous EtOH (11 mL). This was followed by the portion wise addition of sodium (782 mg, 23.0 mmol, 1 equiv.) and the resulting mixture was then allowed to stir vigorously until a homogeneous solution formed. After cooling to $-10\text{ }^{\circ}\text{C}$ (acetone/ice), ethyl acetoacetate

(4.34 mL, 34.0 mmol, 1 equiv.) was added dropwise to the solution and the reaction mixture was then allowed to warm to rt and stirred for 1 h before being cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) again. A solution of 2,3 dibromopropene (6.80 g, 34.0 mmol, 1 equiv.) in diethyl ether (5 mL) was added dropwise to the reaction mixture, which was then allowed to warm to rt and stirred for 24 h, during which time a brown suspension formed. The suspension was filtered through a plug of celite and the solvent removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 2 – 10% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **17** as a clear oil (7.18 g, 28.8 mmol, 85%).

IR, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.⁵³

Rf: 0.44 (10% EtOAc in PE); ^1H NMR (400 MHz, DMSO- d_6) δ 1.19 (t, $J = 7.2$ Hz, 3H, OCH_2CH_3), 2.23 (s, 3H, CHCOCH_3), 2.95 – 2.82 (m, 2H, BrCCH_2CH), 3.95 (t, $J = 7.2$ Hz, 1H, CH_2CHCO), 4.14 (q, $J = 7.2$ Hz, 2H, OCH_2CH_3), 5.49 (d, $J = 2.0$ Hz, 1H, $\text{CH}_2=\text{CBr}$), 5.75 (d, $J = 2.0$ Hz, 1H, $\text{CH}_2=\text{CBr}$) ppm; ^{13}C NMR (101 MHz, DMSO- d_6) δ 13.9, 29.1, 38.8, 57.3, 61.2, 119.6, 129.7, 168.0, 201.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_9\text{H}_{13}\text{O}_3\text{Br}$: 249.0126; found: 249.0122.

5-Bromohex-5-en-2-one (19)²⁴

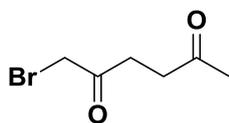
A flame-dried, 2-neck round-bottomed flask was purged with Ar and to this was added anhydrous EtOH (32 mL). This was followed by the portion-wise addition of sodium (1.84 g, 80.0 mmol, 1 equiv.) and the resulting mixture was then allowed to stir vigorously until a homogeneous

solution formed. After cooling to $-10\text{ }^{\circ}\text{C}$ (acetone/ice), ethyl acetoacetate (11.2 mL, 88.0 mmol, 1.1 equiv.) was added dropwise to the solution and the reaction mixture was then allowed to warm to rt and stirred for 1 h before being cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) again. This was followed by the dropwise addition of 2,3 dibromopropene (16.0 g, 80.0 mmol, 1 equiv.) and the reaction mixture was then allowed to stir at $80\text{ }^{\circ}\text{C}$ for 1 h. After cooling to room temperature, 10% NaOH solution (64 mL) was added and the reaction mixture was then allowed to stir at $80\text{ }^{\circ}\text{C}$ for a further 2 h, during which time the solution turned orange in colour. After cooling to room temperature, the reaction mixture was then acidified to $\text{pH} = 3$ with concentrated HCl and allowed to stir at $80\text{ }^{\circ}\text{C}$ for a further 2 h. After cooling to room temperature, the aqueous layer was extracted with aliquots of diethyl ether (3×50 mL) and the combined organic layers were then washed with a saturated solution of NaHCO_3 (30 mL), brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% Et_2O in PE. Pure fractions were evaporated to dryness to afford compound **19** as a clear liquid (8.36 g, 47.2 mmol, 59%).

^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.^{24, 54}

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Rf: 0.25 (10% Et₂O in PE); ¹H NMR (300 MHz, CDCl₃) δ 2.11 (s, 3H, CH₂COCH₃), 2.66 – 2.62 (m, 4H, CCH₂CH₂C and CCH₂CH₂C), 5.34 (d, *J* = 1.8 Hz, 1H, CH₂=CBr), 5.55 (d, *J* = 1.8 Hz, 1H, CH₂=CBr) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 30.1, 35.4, 41.9, 117.5, 132.8, 206.5 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₆H₁₀OBr: 176.9915; found: 176.9913.

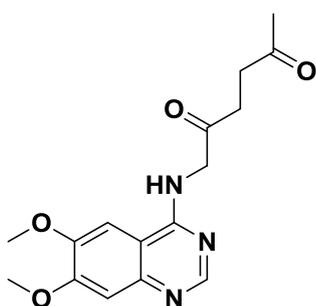
1-Bromohexane-2,5-dione (20)²⁵

A 2-neck round-bottom flask was charged with 5-bromohex-5-en-2-one (300 mg, 1.69 mmol, 1 equiv.), purged with Ar and dissolved in a 4:1 solution of acetonitrile/H₂O (10 mL).

To this was added *N*-bromosuccinimide (391 mg, 2.20 mmol, 1.3 equiv.) in one portion,

followed by the addition of a catalytic amount of hydrobromic acid. The resulting yellow solution was then allowed to stir at rt for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was diluted with diethyl ether (30 mL), treated with a sufficient amount of 10 mol% sodium thiosulphate so that the solution became clear and the organic layer was then separated. The organic layer was then washed with a saturated solution of NaHCO₃ (20 mL), brine (20 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **20** as a clear oil (245 mg, 1.27 mmol, 75%).

Rf: 0.22 (10% EtOAc in PE); ¹H NMR (300 MHz, DMSO-*d*₆) δ 2.17 (s, 3H, CH₂COCH₃), 2.80 – 2.74 (m, 2H, CCH₂CH₂C), 2.88 – 2.82 (m, 2H, CCH₂CH₂C), 3.97 (s, 2H, BrCH₂CO) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 29.8, 33.5, 34.5, 37.6, 201.1, 206.7 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₆H₁₀O₂Br: 192.9864; found: 192.9869.

1-[(6,7-Dimethoxyquinazolin-4-yl)amino]hexane-2,5-dione (21)²⁶

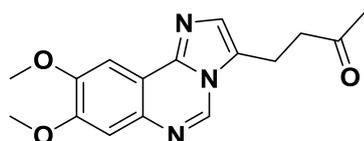
A 2-neck round-bottom flask was charged with 6,7-dimethoxyquinazolin-4-amine (530 mg, 2.58 mmol, 1 equiv.), purged with Ar and dissolved in anhydrous EtOH (20 mL). To this was added a solution of 1-bromohexane-2,5-dione (619 mg, 3.21 mmol, 1.25 equiv.) in anhydrous EtOH (5 mL) and the reaction mixture was then

allowed to stir at 80 °C for 6 h during which time a precipitate formed. After cooling to rt, the solvent was removed *in vacuo* and the residue was diluted with H₂O (10 mL), 1 M NaOH solution (20 mL) and DCM (50 mL) and the organic layer was

separated. The aqueous layer was extracted with aliquots of DCM (3 × 20 mL) and the combined organic layers were then washed with a saturated solution of NaHCO₃ (30 mL), brine (30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 2 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **21** as a tan solid (168 mg, 1.36 mmol, 53%).

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Rf: 0.22 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3303 (m, N-H stretch), 1724 (s, C=O stretch), 1701 (s, C=O stretch), 1621 (m, N-H bend), 1507 (m, C-H bend), 1287 (s, C-N stretch), 1219 (s, C-O-C stretch), 772 (s, N-H wag); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 2.10 (s, 3H, CH_2COCH_3), 2.93 – 2.75 (m, 4H, $\text{CCH}_2\text{CH}_2\text{C}$ and $\text{CCH}_2\text{CH}_2\text{C}$), 3.94 (s, 3H, ArOCH_3), 4.04 (s, 3H, ArOCH_3), 5.81 (s, 2H, ArNHCH_2CO), 7.10 (s, 1H, ArH), 8.00 (s, 1H, ArH), 8.68 (s, 1H, ArH), 9.57 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 29.5, 33.4, 36.8, 56.9, 57.1, 58.8, 98.8, 105.1, 105.8, 135.5, 149.7, 152.5, 156.7, 161.5, 202.3, 207.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{16}\text{H}_{20}\text{N}_3\text{O}_4$: 318.1456; found: 318.1454.

4-(8,9-Dimethoxyimidazo[1,2-c]quinazolin-3-yl)butan-2-one (22)

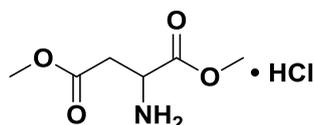
Compound **22** was prepared in the same manner as the as the previous compound **21**. 6,7-Dimethoxyquinazolin-4-amine (530 mg, 2.58 mmol, 1 equiv.) was coupled and purification was achieved using flash column

chromatography with elution gradient of 2 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **22** as an off-white solid (154 mg, 0.514 mmol, 20%).

Rf: 0.44 (10% MeOH in DCM); IR (ATR) cm^{-1} : 1705 (s, C=O stretch), 1497 (m, C-H bend), 1223 (s, C-O-C stretch); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 2.15 (s, 3H, CH_2COCH_3), 2.91 (br s, 4H, $\text{CCH}_2\text{CH}_2\text{C}$ and $\text{CCH}_2\text{CH}_2\text{C}$), 3.91 (s, 3H, ArOCH_3), 3.95 (s, 3H, ArOCH_3), 7.34 (s, 1H, ArH), 7.64 (s, 1H, ArH), 7.70 (s, 1H, ArH), 9.08 (s, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 22.5, 29.8, 42.0, 55.8, 55.8, 101.6, 108.7, 109.0, 112.4, 135.8, 141.4, 145.1, 149.8, 151.1, 207.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{16}\text{H}_{18}\text{N}_3\text{O}_3$: 300.1348; found: 300.1347.

General procedure for aspartic acid esterification:⁵⁵

A 2-neck round-bottom flask was charged with aspartic acid (1 equiv.) and to this was added the respective alcohol for esterification (2 mL/mmol aspartic acid). The resulting suspension was cooled to $-10\text{ }^\circ\text{C}$ (acetone/ice) and this was followed by the dropwise addition of thionyl chloride (1.4 equiv.) over 15 min. The reaction mixture was then allowed to stir at reflux for 6 h and the excess thionyl chloride was then removed by vacuum distillation. The oily residue was triturated with diethyl ether and the white solid that formed was collected by filtration and dried under vacuum to afford the respective compound as the hydrochloride salt with no further purification required.

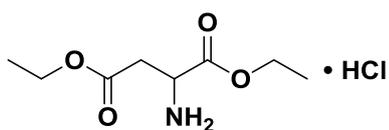
Dimethyl aspartate hydrochloride (24)

Compound **24** was prepared according to the general procedure for aspartic acid esterification. Aspartic acid (4.00 g, 27.2 mmol) was esterified with MeOH to afford compound **24** as a white solid (5.10 g, 25.8 mmol, 95%).

^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{55, 56}

Mp: 115 – 117 $^\circ\text{C}$; ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 3.00 (d, $J = 5.4$ Hz, 2H, COCH_2CH), 3.63 (s, 3H, COOCH_3), 3.71 (s, 3H, COOCH_3), 4.24 (t, $J = 5.5$ Hz, 1H, CH_2CHNH_2), 8.32 (br s, 3H, $\text{CHNH}_2\cdot\text{HCl}$) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 34.11, 48.47, 52.12, 52.99, 168.82, 169.62 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_6\text{H}_{12}\text{NO}_4$: 162.0766; found: 162.0769.

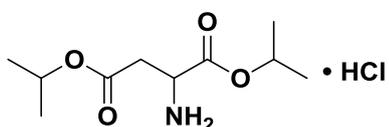
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Diethyl aspartate hydrochloride (25)

Compound **25** was prepared according to the general procedure for aspartic acid esterification. Aspartic acid (4.00 g, 27.2 mmol) was esterified with EtOH to afford compound **25** as a white solid (6.01 g, 26.6 mmol, 98%).

^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.^{57, 58}

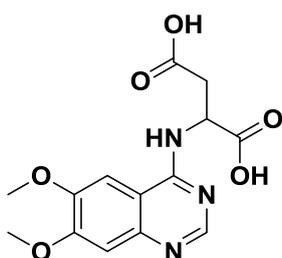
mp: 108 – 110 °C; ^1H NMR (300 MHz, DMSO- d_6) δ 1.19 (2 \times t, J = 7.2, 6H, 2 \times OCH₂CH₃), 3.04 – 2.87 (m, 2H, COCH₂CH), 4.23 – 4.05 (m, 5H, CH₂CHNH₂ and 2 \times OCH₂CH₃), 8.29 (br s, 3H, CHNH₂·HCl) ppm; ^{13}C NMR (75 MHz, DMSO- d_6) δ 13.9, 14.0, 34.7, 48.7, 60.8, 61.8, 168.8, 169.2 ppm; HRMS-TOF MS ESI+: m/z [M+H]⁺ calculated for C₈H₁₆NO₄: 190.1079; found: 190.1077.

Diisopropyl aspartate hydrochloride (26)

Compound **26** was prepared according to the general procedure for aspartic acid esterification. Aspartic acid (4.00 g, 27.2 mmol) was esterified with IPA to afford compound **26** as a white solid (6.63 g, 26.1 mmol, 96%).

^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.⁵⁹

^1H NMR (300 MHz, DMSO- d_6) δ 1.26 – 1.44 (m, 12H, 2 \times (CH₃)₂CHO), 3.03 – 2.83 (m, 2H, CH₂CHNH₂), 4.18 (t, J = 5.2 Hz, 1H, CH₂CHNH₂), 5.04 – 4.84 (m, 2H, 2 \times (CH₃)₂CHO), 8.54 (s, 3H, CH₂CHNH₂·HCl) ppm; ^{13}C NMR (75 MHz, DMSO- d_6) δ 21.3, 21.4, 21.5, 21.5, 34.6, 48.6, 68.5, 69.8, 168.0, 168.7 ppm; HRMS-TOF MS ESI+: m/z [M+H]⁺ calculated for C₁₀H₂₀NO₄: 218.1392; found: 218.1388.

(6,7-Dimethoxyquinazolin-4-yl)aspartic acid (27)

A microwave vial was charged with aspartic acid (272 mg, 2.04 mmol, 1 equiv.), 4-chloro-6,7-dimethoxyquinazoline (550 mg, 2.45 mmol, 1.2 equiv.), purged with Ar and to this was added a 1:1 solution of H₂O/DMSO (10 mL). Et₃N (1.14 mL, 8.16 mmol, 4 equiv.) was then added and after capping the vial the resulting suspension was warmed to 80 °C and allowed to stir for 72 h. After allowing to cool to rt, the solvent was removed by freeze drying and the residue was then diluted with DCM (30 mL), filtered and

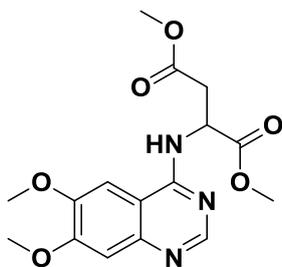
washed with aliquots of DCM (3 \times 10 mL) and cold MeOH (10 mL). The collected solid was dried under vacuum to afford compound **27** as a white solid (590 mg, 1.84 mmol, 90%) with no further purification required.

Rf: 0.10 (30% MeOH in DCM); IR (ATR) cm⁻¹: 3141 (m, N-H stretch), 2971 (s, O-H stretch), 1765 (s, C=O), 1627 (m, N-H bend), 1509 (m, C-H bend), 1277 (s, C-N stretch), 1247 (m, C-O-C stretch), 770 (s, N-H wag); ^1H NMR (300 MHz, DMSO- d_6) δ 2.89 (ddd, J = 24.5, 16.5, 7.0 Hz, 2H, CH₂CHNH), 3.90 (s, 6H, 2 \times ArOCH₃), 5.14 (d, J = 5.3 Hz, 1H, CH₂CHNH), 7.12 (s, 1H, ArH), 7.64 (s, 1H, ArH), 8.17 (d, J = 6.3 Hz, 1H, ArNH), 8.37 (s, 1H, ArH) ppm; ^{13}C NMR (75 MHz, DMSO- d_6) δ 36.1, 50.4, 55.8, 56.1, 102.0, 106.7, 108.4, 145.7, 148.6, 152.9, 154.1, 157.9, 171.9, 173.1 ppm; HRMS-TOF MS ESI+: m/z [M+H]⁺ calculated for C₁₄H₁₆N₃O₆: 322.1039; found: 322.1046.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

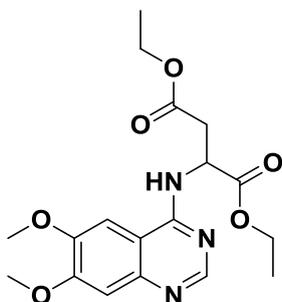
General procedure for aspartic derivative coupling:

A microwave vial was charged with 4-chloro-6,7-dimethoxyquinazoline (1 equiv.), aspartic derivative (1.1 equiv.), purged with Ar and to this was added acetonitrile (5 mL/mmol 4-chloro-6,7-dimethoxyquinazoline). After the addition of a catalytic amount of 4 M HCl in dioxane the vial was capped, and the resulting suspension was warmed to 100 °C and allowed to stir for 24 h. After allowing to cool to room temperature, the reaction mixture was diluted with a saturated solution of NaHCO₃ (10 mL/mmol 4-chloro-6,7-dimethoxyquinazoline), H₂O (10 mL/mmol 4-chloro-6,7-dimethoxyquinazoline) and DCM (10 mL/mmol 4-chloro-6,7-dimethoxyquinazoline) and the organic layer was separated. The aqueous layer was then extracted with aliquots of DCM (3 × 10 mL/mmol 4-chloro-6,7-dimethoxyquinazoline) and the combined organic layers were then washed with a saturated solution of brine (10 mL/mmol 4-chloro-6,7-dimethoxyquinazoline), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography. Pure fractions were evaporated to dryness to afford the respective compound.

Dimethyl (6,7-dimethoxyquinazolin-4-yl)aspartate (28)

Compound **28** was prepared according to the general procedure for aspartic derivative coupling. 4-Chloro-6,7-dimethoxyquinazoline (250 mg, 1.11 mmol) was coupled with dimethyl aspartate hydrochloride (242 mg, 1.22 mmol) and purification was achieved using flash column chromatography with elution gradient of 2 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **28** as a yellow solid (287 mg, 0.821 mmol, 74%).

Rf: 0.39 (5% MeOH in DCM); IR (ATR) cm⁻¹: 3283 (m, N-H stretch), 1732 (s, C=O stretch), 1616 (m, N-H bend), 1506 (m, C-H bend), 1289 (s, C-N stretch), 1219 (s, C-O-C stretch), 1170 (s, C-O stretching), 844 (s, N-H wag); ¹H NMR (300 MHz, DMSO-*d*₆) δ 3.19 (d, *J* = 6.9 Hz, 2H, CH₂CHNH), 3.63 (s, 3H, COOCH₃), 3.67 (s, 3H, COOCH₃), 3.97 (s, 3H, ArOCH₃), 3.98 (s, 3H, ArOCH₃), 5.42 (td [app. q], *J* = 7.0 Hz, 1H, CH₂CHNH), 7.40 (s, 1H, ArH), 8.26 (s, 1H, ArH), 8.87 (s, 1H, ArH), 10.62 (d, *J* = 7.4 Hz, 1H, ArNH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 34.8, 50.9, 51.9, 52.7, 56.4, 57.0, 99.5, 104.0, 106.6, 134.6, 148.5, 150.2, 156.3, 159.1, 170.1, 170.4 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₆H₂₀N₃O₆: 350.1352; found: 350.1350.

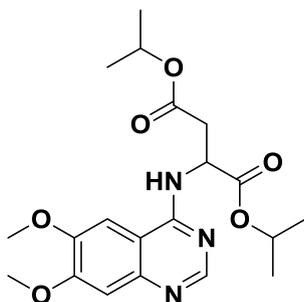
Diethyl (6,7-dimethoxyquinazolin-4-yl)aspartate (29)

Compound **29** was prepared according to the general procedure for aspartic derivative coupling. 4-Chloro-6,7-dimethoxyquinazoline (250 mg, 1.11 mmol) was coupled with diethyl aspartate hydrochloride (275 mg, 1.22 mmol) and purification was achieved using flash column chromatography with elution gradient of 2 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **29** as a yellow-brown solid (331 mg, 0.877 mmol, 79%).

Rf: 0.35 (5% MeOH in DCM); IR (ATR) cm⁻¹: 3233 (m, N-H stretch), 1732 (s, C=O stretch), 1616 (m, N-H bend), 1501 (m, C-H bend), 1248 (s, C-N stretch), 1211 (s, C-O-C stretch), 1167 (s, C-O stretching), 838 (s, N-H wag);

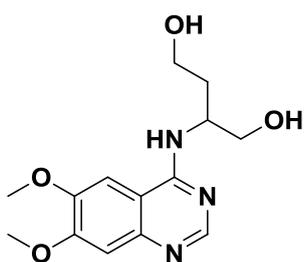
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 1.15 ($2 \times t$, $J = 7.0$ Hz, 6H, $2 \times \text{OCH}_2\text{CH}_3$), 3.16 (d, $J = 7.0$ Hz, 2H, CH_2CHNH), 3.96 (s, 3H, ArOCH_3), 3.98 (s, 3H, ArOCH_3), 4.15 – 4.07 (m, 4H, OCH_2CH_3), 5.38 (td [app. q], $J = 7.0$ Hz, 1H, CH_2CHNH), 7.42 (s, 1H, ArH), 8.28 (s, 1H, ArH), 8.86 (s, 1H, ArH), 10.63 (d, $J = 7.4$ Hz, 1H, ArNH) ppm; ^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 13.9, 14.0, 35.1, 51.1, 56.4, 57.1, 60.5, 61.4, 99.5, 104.1, 106.5, 134.5, 148.5, 150.2, 156.3, 159.2, 169.6, 169.8 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{18}\text{H}_{24}\text{N}_3\text{O}_6$: 378.1665; found: 378.1666.

Diisopropyl (6,7-dimethoxyquinazolin-4-yl)aspartate (30)

Compound **30** was prepared according to the general procedure for aspartic derivative coupling. 4-Chloro-6,7-dimethoxyquinazoline (250 mg, 1.11 mmol) was coupled with diisopropyl aspartate hydrochloride (310 mg, 1.22 mmol) and purification was achieved using flash column chromatography with elution gradient of 2 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **30** as a yellow solid (373 mg, 0.921 mmol, 83%).

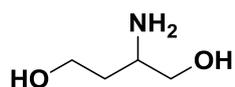
Rf: 0.36 (5% MeOH in DCM); IR (ATR) cm^{-1} : 3233 (m, N-H stretch), 1730 (s, C=O stretch), 1618 (m, N-H bend), 1505 (m, C-H bend), 1250 (s, C-N stretch), 1213 (s, C-O-C stretch), 1170 (s, C-O stretching), 840 (s, N-H wag); ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 1.16 ($2 \times d$, $J = 6.2$ Hz, 12H, $2 \times (\text{CH}_3)_2\text{CHO}$), 3.13 – 3.07 (m, 2H, CH_2CHNH), 3.96 (s, 3H, ArOCH_3), 3.98 (s, 3H, ArOCH_3), 4.98 – 48.7 (m, 2H, $2 \times (\text{CH}_3)_2\text{CHO}$), 5.30 (td [app. q], $J = 7.2$ Hz, 1H, CH_2CHNH), 7.41 (s, 1H, ArH), 8.26 (s, 1H, ArH), 8.86 (s, 1H, ArH), 10.56 (d, $J = 7.2$ Hz, 1H, ArNH) ppm; ^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 21.4, 21.4, 21.5, 21.5, 35.5, 51.3, 56.4, 57.0, 67.9, 69.1, 99.6, 104.0, 106.5, 134.6, 148.5, 150.2, 156.2, 159.1, 169.0, 169.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{20}\text{H}_{28}\text{N}_3\text{O}_6$: 406.1978; found: 406.1977.

2-[(6,7-Dimethoxyquinazolin-4-yl)amino]butane-1,4-diol (32)

Compound **32** was prepared according to the general procedure for aspartic derivative coupling. 4-Chloro-6,7-dimethoxyquinazoline (712 mg, 3.17 mmol) was coupled with 2-aminobutane-1,4-diol (367 mg, 3.49 mmol) and purification was achieved using flash column chromatography with elution gradient of 10 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **32** as a yellow solid (772 mg, 2.631 mmol, 83%).

Rf: 0.29 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3359 (s, O-H stretch), 3279 (m, N-H stretch), 3001 (w, O-H stretch), 1623 (m, N-H bend), 1505 (m, C-H bend), 1250 (s, C-N stretch), 1215 (s, C-O-C stretch), 838 (s, N-H wag); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 1.83 – 1.69 (m, 1H, $\text{OHCH}_2\text{CH}_2\text{CH}$), 2.00 – 1.86 (m, 1H, $\text{OHCH}_2\text{CH}_2\text{CH}$), 3.63 – 3.45 (m, 4H, $2 \times \text{CH}_2\text{OH}$), 3.88 (s, 3H, ArOCH_3), 3.90 (s, 3H, ArOCH_3), 4.50 – 4.44 (m, 1H, CH_2CHNH), 4.81 (br s, 1H, CH_2OH), 4.55 (br s, 1H, CH_2OH), 7.07 (s, 1H, ArH), 7.50 (d, $J = 8.0$ Hz, 1H, ArNH), 7.66 (s, 1H, ArH), 8.30 (s, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 34.2, 49.7, 55.7, 56.2, 58.1, 63.1, 102.3, 107.0, 108.5, 146.0, 148.2, 153.4, 153.7, 158.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{14}\text{H}_{20}\text{N}_3\text{O}_4$: 294.1454; found: 294.1461.

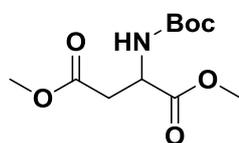
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

2-Aminobutane-1,4-diol (33)²⁹

A flame-dried, 2-neck round-bottomed flask was charged with lithium aluminum hydride (1.20 g, 31.7 mmol, 3 equiv.), purged with Ar and to this was added THF (20 mL). The resulting grey slurry was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) and a solution of diethyl aspartate (2.00 g, 10.6 mmol, 1 equiv.) in THF (10 mL) was added dropwise over 15 min during which time vigorous bubbling occurred. The resulting solution was then allowed to stir at $70\text{ }^{\circ}\text{C}$ for 1 h and after allowing to cool to rt was then cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice). To this solution was added 2-propanol (15 mL), followed by the addition of H_2O (3.71 mL, 206 mmol, 6.5 equiv. [to LiAlH_4]) and the resulting slurry was then allowed to stir vigorously for 15 min. After removal of the solvent *in vacuo*, the dry aluminum salts were placed in a Soxhlet cup and continuously extracted with 2-propanol for 12 h via Soxhlet extraction. After removal of the solvent *in vacuo*, purification of the crude product was achieved by kugelrohr distillation. Fractions with boiling point between $124\text{ }^{\circ}\text{C}$ – $126\text{ }^{\circ}\text{C}$ at ~ 20 mbar were collected to afford compound **33** as a light-yellow oil (914 mg, 8.70 mmol, 82%).

IR, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.²⁹

^1H NMR (300 MHz, CD_3OD) δ 1.57 – 1.43 (m, 1H, $\text{OHCH}_2\text{CH}_2\text{CH}$), 1.74 – 1.61 (m, 1H, $\text{OHCH}_2\text{CH}_2\text{CH}$), 2.97 – 2.88 (m, 1H, CH_2CHCH_2), 3.40 – 3.32 (m, 1H, CHCH_2OH), 3.56 – 3.49 (m, 1H, CHCH_2OH), 3.68 (m [app t.], $J = 6.4$ Hz, 2H, $\text{OHCH}_2\text{CH}_2\text{CH}$) ppm; ^{13}C NMR (75 MHz, CD_3OD) δ 37.0, 51.7, 60.5, 67.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_4\text{H}_{12}\text{NO}_2$: 106.0868; found: 106.0872.

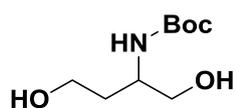
Dimethyl (tert-butoxycarbonyl)aspartate (42)³⁸

A 2-neck round-bottom flask was charged with dimethyl aspartate hydrochloride (1.95 g, 9.87 mmol, 1 equiv.), purged with Ar and dissolved in anhydrous MeOH (20 mL). After the addition of Et_3N (3.03 mL, 21.7 mmol, 2.2 equiv.), the solution was cooled to $0\text{ }^{\circ}\text{C}$ and a solution of Boc_2O (2.37, 10.8 mmol, 1.1 equiv.) in anhydrous MeOH (10 mL) was added slowly. The reaction was then allowed to warm to rt and stirred for 16 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 20 – 40% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **42** as a white solid (2.22 g, 8.49 mmol, 86%).

Mp, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{38, 60}

Rf: 0.35 (30% EtOAc in PE); mp: $64 - 66\text{ }^{\circ}\text{C}$; ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 1.37 (s, 9H, $(\text{CH}_3)_3\text{OCO}$), 2.82 – 2.62 (m, 2H, COCH_2CH), 3.60 (s, 3H, COOCH_3), 3.63 (s, 3H, COOCH_3), 4.45 – 4.31 (m, $J = 13.8, 7.6$ Hz, 1H, NHCHCH_2), 7.27 (d, $J = 8.2$ Hz, 1H, CONHCH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 28.1, 35.7, 50.1, 51.6, 52.1, 78.5, 155.2, 170.5, 171.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{Na}]^+$ calculated for $\text{C}_{11}\text{H}_{19}\text{NO}_6\text{Na}$: 284.1105; found: 284.1113.

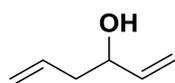
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

tert-Butyl (1,4-dihydroxybutan-2-yl)carbamate (40)²⁹

A flame-dried, 2-neck round-bottomed flask was charged with lithium aluminum hydride (871 mg, 23.0 mmol, 3 equiv.), purged with Ar and to this was added THF (20 mL). The resulting grey slurry was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) and a solution of dimethyl (*tert*-butoxycarbonyl)aspartate (2.00 g, 7.65 mmol, 1 equiv.) in THF (10 mL) was added dropwise over 15 min during which vigorous bubbling occurred. The resulting solution was then allowed to stir at $70\text{ }^{\circ}\text{C}$ for 1 h and after allowing to cool to rt was then cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice). To this solution was added 2-propanol (15 mL), followed by the addition of H_2O (2.69 mL, 149 mmol, 6.5 equiv. [to LiAlH_4]) and the resulting slurry was then allowed to stir vigorously for 15 min. After removal of the solvent *in vacuo*, the dry aluminum salts were placed in a Soxhlet cup and continuously extracted with 2-propanol for 12 h via Soxhlet extraction. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **40** as a white solid (1.27 g, 6.20 mmol, 81%).

Mp, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{60, 61}

Rf: 0.29 (10% MeOH in DCM); mp: 62 – 63 $^{\circ}\text{C}$; ^1H NMR (300 MHz, CDCl_3) δ 1.28 (s, 9H, $(\text{CH}_3)_3\text{OCO}$), 1.53 – 1.39 (m, 1H, $\text{CH}_2\text{CH}_2\text{CH}$), 1.70 – 1.56 (m, 1H, 1H, $\text{CH}_2\text{CH}_2\text{CH}$), 3.61 – 3.38 (m, 5H, $2 \times \text{CH}_2\text{OH}$ and CH_2CHNH), 4.30 (t, $J = 5.0\text{ Hz}$, 1H, CH_2OH), 4.42 (t, $J = 5.0\text{ Hz}$, 1H, CH_2OH), 5.44 (d, $J = 8.2\text{ Hz}$, 1H, CONHCH), ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 28.2, 34.4, 49.4, 58.5, 64.4, 79.4, 156.8 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{Na}]^+$ calculated for $\text{C}_9\text{H}_{19}\text{NO}_4\text{Na}$: 228.1207; found: 228.1201.

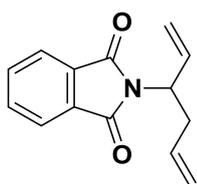
Hexa-1,5-dien-3-ol (45)³⁹

A flame-dried, 2-neck round-bottomed flask was charged with oven-dried magnesium turnings (8.51 g, 350 mmol, 1.4 equiv.) and purged with Ar. To this was added of diethyl ether (25 mL), followed by the dropwise addition of allyl bromide (4.32 mL, 50.0 mmol, 20 mol%) to initiate the Grignard reaction. This was followed by the dropwise addition of a solution of allyl bromide (21.6 mL, 250 mmol, 1 equiv.) and acrolein (16.7 mL, 250 mmol, 1 equiv.) in diethyl ether (75 mL) over 3 h whilst maintaining a gentle reflux. The reaction mixture was then allowed to stir at $45\text{ }^{\circ}\text{C}$ for 3 h and cooled to $0\text{ }^{\circ}\text{C}$. A saturated solution of NH_4Cl (100 mL) was added dropwise resulting in an exotherm, followed by the addition of H_2O (125 mL) and the resulting suspension was then allowed to stir at rt for 16 h. The reaction mixture was acidified to $\text{pH} = 2$ with a solution of 2 M HCl and the aqueous layer was then extracted with aliquots of diethyl ether ($3 \times 100\text{ mL}$). The combined organic layers were then washed with a saturated solution brine (50 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved by kugelrohr distillation. Fractions with boiling point between $65\text{ }^{\circ}\text{C} - 70\text{ }^{\circ}\text{C}$ at $\sim 60\text{ mbar}$ were collected to afford compound **45** as a clear oil (14.1 g, 143 mmol, 57%).

IR, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{39, 62}

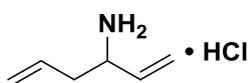
Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Rf: 0.38 (15% EtOAc in PE); ^1H NMR (300 MHz, CDCl_3) δ 2.34 – 2.24 (m, 2H, CHCH_2CH), 4.18 – 4.08 (m, 1H, CH_2CHCH), 5.09 (dd, $J = 10.3, 1.2$ Hz, 2H, $\text{CH}=\text{CH}_2$), 5.14 – 5.10 (m, 1H, $\text{CH}=\text{CH}_2$), 5.21 (dt, $J = 16.2, 1.2$ Hz, 1H, $\text{CH}_2\text{CH}=\text{CH}_2$), 5.92 – 5.69 (m, 2H, $\text{CH}=\text{CH}_2$ and $\text{CH}=\text{CH}_2$) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 41.7, 71.9, 114.8, 118.2, 134.2, 140.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_6\text{H}_{11}\text{O}$: 98.0731; found: 98.0732.

2-(Hexa-1,5-dien-3-yl)isoindoline-1,3-dione (46)⁶³

A flame-dried, 2-neck round-bottomed flask was charged with hexa-1,5-dien-3-ol (1.20 g, 12.2 mmol, 1 equiv.), purged with Ar and dissolved in THF (100 mL). After cooling the solution to 0 °C, triphenylphosphine (3.53 g, 13.4 mmol, 1.1 equiv.) and phthalimide (1.98 g, 13.4 mmol, 1.1 equiv.) were added as a single portion and this was followed by the dropwise addition of a solution of di-*tert*-butyl azodicarboxylate (3.66 g, 15.9 mmol, 1.3 equiv.) in THF (20 mL). The solution was then allowed to warm to rt and stirred for 16 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, the residue was triturated with a 1:1 solution of diethyl ether/PE (60 mL) which resulted in the precipitation of triphenylphosphine oxide. The slurry was then filtered through a plug of celite and the solvent removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 20% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **46** as a clear oil (2.11 g, 9.29 mmol, 76%).

Rf: 0.67 (20% MeOH in DCM); ^1H NMR (300 MHz, CDCl_3) δ 2.65 – 2.53 (m, 1H, NCHCH_2), 2.93 – 2.77 (m, 1H, NCHCH_2), 4.86 – 4.73 (m, 1H, NCHCH_2), 4.93 (d, $J = 10.1$ Hz, 1H, $\text{CH}=\text{CH}_2$), 5.01 (dd, $J = 17.2, 1.0$ Hz, 1H, $\text{CH}=\text{CH}_2$), 5.15 (d, $J = 10.2$ Hz, 1H, $\text{CH}=\text{CH}_2$), 5.20 (d, $J = 17.2$ Hz, 1H, $\text{CH}=\text{CH}_2$), 5.76 – 5.57 (m, 1H, $\text{CH}=\text{CH}_2$), 6.25 – 6.09 (m, 1H, $\text{CH}=\text{CH}_2$), 7.71 – 7.61 (m, 2H, $2 \times \text{ArH}$), 7.82 – 7.72 (m, 2H, $2 \times \text{ArH}$) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 36.5, 53.4, 117.4, 118.2, 123.2, 131.9, 133.9, 134.0, 135.3, 168.0 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{14}\text{H}_{14}\text{NO}_2$: 228.1025; found: 228.1019.

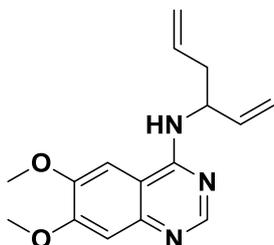
Hexa-1,5-dien-3-amine hydrochloride (48)⁶³

A 2-neck round-bottomed flask was charged with 2-(hexa-1,5-dien-3-yl)isoindoline-1,3-dione (1.80 g, 7.92 mmol, 1 equiv.), purged with Ar and dissolved in EtOH (40 mL). To the resulting solution was added hydrazine hydrate (0.58 mL, 11.9 mmol, 1.5 equiv.) and the solution was then allowed to stir at 80 °C for 6 h or until complete consumption of the starting material as indicated by TLC. After cooling to room temperature, the reaction mixture was then acidified to pH = 2 with concentrated HCl and allowed to stir at rt for 1 h. The slurry was then filtered through a plug of celite and the solvent removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 10 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **48** as a white solid (762 mg, 5.70 mmol, 72%).

^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.⁶⁴

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

Rf: 0.28 (10% MeOH in DCM); ^1H NMR (300 MHz, CD_3OD) δ 1.01 – 0.85 (m, 2H, NH_2CHCH_2), 3.82 – 3.66 (m, 1H, NH_2CHCH_2), 4.95 (s, 3H, $\text{CHNH}_2\cdot\text{HCl}$), 5.24 – 5.11 (m, 2H, $\text{CH}=\text{CH}_2$), 5.41 – 5.31 (m, 2H, $\text{CH}=\text{CH}_2$), 5.99 – 5.71 (m, 2H, $2 \times \text{CH}=\text{CH}_2$) ppm; ^{13}C NMR (75 MHz, CD_3OD) δ 38.6, 54.4, 120.4, 120.7, 133.0, 134.9 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_6\text{H}_{11}\text{N}$: 98.0970; found: 98.0967.

N-(Hexa-1,5-dien-3-yl)-6,7-dimethoxyquinazolin-4-amine (49)

Compound **49** was prepared according to the general procedure for aspartic derivative coupling. 4-Chloro-6,7-dimethoxyquinazoline (170 mg, 0.757 mmol) was coupled with hexa-1,5-dien-3-amine hydrochloride (112 mg, 0.851 mmol) and purification was achieved using flash column chromatography with elution gradient of 2 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **49** as an off-white solid (160 mg, 0.561 mmol, 74%).

Rf: 0.45 (5% MeOH in DCM); IR (ATR) cm^{-1} : 3283 (m, N-H stretch), 3073 (m, =CH stretch), 1648 (s, C=C stretch), 1621 (m, N-H bend), 1503 (m, C-H bend), 1242 (s, C-N stretch), 1217 (s, C-O-C stretch), 986 (m, C-H bend), 912 (m, C-H bend), 848 (s, N-H wag); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 2.56 – 2.51 (m, 2H, NHCHCH_2), 3.89 (s, 3H, ArOCH_3), 3.90 (s, 3H, ArOCH_3), 5.28 – 5.02 (m, 5H, $4 \times \text{CH}=\text{CH}_2$ and NHCHCH_2), 5.83 (ddt, $J = 17.2, 10.3, 6.8$ Hz, 1H, $\text{CH}_2\text{CH}=\text{CH}_2$), 5.99 (ddd, $J = 17.2, 10.3, 5.8$ Hz, 1H, $\text{CHCH}=\text{CH}_2$), 7.09 (s, 1H, ArH), 7.68 (s, 1H, ArH), 7.74 (d, $J = 8.2$ Hz, 1H, ArNH), 8.32 (s, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 38.4, 51.8, 55.7, 56.2, 102.0, 107.1, 108.4, 114.6, 117.0, 135.4, 138.9, 146.2, 148.3, 153.4, 153.8, 157.8 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{16}\text{H}_{20}\text{N}_3\text{O}_2$: 286.1556; found: 286.1559.

5.9 References

1. T. Barf and A. Kaptein, *Journal of Medicinal Chemistry*, 2012, **55**, 6243-6262.
2. J. Beenstock, N. Mooshayef and D. Engelberg, *Trends in Biochemical Sciences*, 2016, **41**, 938-953.
3. B. Nolen, S. Taylor and G. Ghosh, *Molecular Cell*, 2004, **15**, 661-675.
4. T. Grabe, J. Lategahn and D. Rauh, *ACS Medicinal Chemistry Letters*, 2018, **9**, 779-782.
5. A. C. Carrera, K. Alexandrov and T. M. Roberts, *Proceedings of the National Academy of Sciences of the United States of America*, 1993, **90**, 442-446.
6. M. J. Robinson, P. C. Harkins, J. Zhang, R. Baer, J. W. Haycock, M. H. Cobb and E. J. Goldsmith, *Biochemistry*, 1996, **35**, 5641-5646.
7. G. H. Iyer, S. Garrod, V. L. Woods and S. S. Taylor, *Journal of Molecular Biology*, 2005, **351**, 1110-1122.
8. E. R. Wood, A. T. Truesdale, O. B. McDonald, D. Yuan, A. Hassell, S. H. Dickerson, B. Ellis, C. Pennisi, E. Horne, K. Lackey, K. J. Alligood, D. W. Rusnak, T. M. Gilmer and L. Shewchuk, *Cancer Research*, 2004, **64**, 6652-6659.
9. Ketan S. Gajiwala, J. Feng, R. Ferre, K. Ryan, O. Brodsky, S. Weinrich, John C. Kath and A. Stewart, *Structure*, 2013, **21**, 209-219.
10. Z. Zhao and P. E. Bourne, *Drug Discovery Today*, 2018, **23**, 727-735.
11. A. Wang, X. Li, H. Wu, F. Zou, X.-E. Yan, C. Chen, C. Hu, K. Yu, W. Wang, P. Zhao, J. Wu, Z. Qi, W. Wang, B. Wang, L. Wang, T. Ren, S. Zhang, C.-H. Yun, J. Liu and Q. Liu, *Journal of Medicinal Chemistry*, 2017, **60**, 2944-2962.
12. J. Pettinger, K. Jones and M. D. Cheeseman, *Angewandte Chemie International Edition*, 2017, **56**, 15200-15209.
13. M. P. Wymann, G. Bulgarelli-Leva, M. J. Zvelebíl, L. Pirola, B. Vanhaesebroeck, M. D. Waterfield and G. Panayotou, *Molecular and Cellular Biology*, 1996, **16**, 1722-1733.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

14. E. H. Walker, M. E. Pacold, O. Perisic, L. Stephens, P. T. Hawkins, M. P. Wymann and R. L. Williams, *Molecular Cell*, 2000, **6**, 909-919.
15. N. T. Ihle, R. Williams, S. Chow, W. Chew, M. I. Berggren, G. Paine-Murrieta, D. J. Minion, R. J. Halter, P. Wipf, R. Abraham, L. Kirkpatrick and G. Powis, *Molecular Cancer Therapeutics*, 2004, **3**, 763-772.
16. I. M. Bell, S. M. Stirdivant, J. Ahern, J. C. Culbertson, P. L. Darke, C. J. Dinsmore, R. A. Drakas, S. N. Gallicchio, S. L. Graham, D. C. Heimbrook, D. L. Hall, J. Hua, N. R. Kett, A. S. Kim, M. Kornienko, L. C. Kuo, S. K. Munshi, A. G. Quigley, J. C. Reid, B. W. Trotter, L. H. Waxman, T. M. Williams and C. B. Zartman, *Biochemistry*, 2005, **44**, 9430-9440.
17. H. P. Morgan, M. J. Walsh, E. A. Blackburn, M. A. Wear, M. B. Boxer, M. Shen, H. Veith, I. W. McNae, M. W. Nowicki, P. A. M. Michels, D. S. Auld, L. A. Fothergill-Gilmore and M. D. Walkinshaw, *Biochemical Journal*, 2012, **448**, 67-72.
18. J. A. Lumley, *QSAR & Combinatorial Science*, 2005, **24**, 1066-1075.
19. D. J. Maly, J. A. Allen and K. M. Shokat, *Journal of the American Chemical Society*, 2004, **126**, 9160-9161.
20. R. Dasari, A. De Carvalho, D. C. Medellin, K. N. Middleton, F. Hague, M. N. M. Volmar, L. V. Frolova, M. F. Rossato, J. J. De La Chapa, N. F. Dybdal-Hargreaves, A. Pillai, V. Mathieu, S. Rogelj, C. B. Gonzales, J. B. Calixto, A. Evidente, M. Gautier, G. Munirathinam, R. Glass, P. Burth, S. C. Pelly, W. A. L. van Otterlo, R. Kiss and A. Kornienko, *ChemMedChem*, 2015, **10**, 2014-2026.
21. E. André, B. Campi, M. Trevisani, J. Ferreira, Â. Malheiros, R. A. Yunes, J. B. Calixto and P. Geppetti, *Biochemical Pharmacology*, 2006, **71**, 1248-1254.
22. D. Wang and F. Gao, *Chemistry Central Journal*, 2013, **7**, 95-110.
23. N. Srimongkolpithak, S. Sundriyal, F. Li, M. Vedadi and M. J. Fuchter, *MedChemComm*, 2014, **5**, 1821-1828.
24. J.-D. Liu, S.-H. Wang, F.-M. Zhang, Y.-Q. Tu and Y.-Q. Zhang, *Synlett*, 2009, **2009**, 3040-3042.
25. H. E. Morton and M. R. Leanna, *Tetrahedron Letters*, 1993, **34**, 4481-4484.
26. S. D. Barchéath, R. I. Tawatao, M. Corr, D. A. Carson and H. B. Cottam, *Journal of Medicinal Chemistry*, 2005, **48**, 6409-6422.
27. N. M. Yoon and Y. S. Gyoung, *Journal of Organic Chemistry*, 1985, **50**, 2443-2450.
28. M. Fieser, *Journal of Chemical Education*, 1995, **72**, A94.
29. J. R. Lakanen, A. E. Pegg and J. K. Coward, *Journal of Medicinal Chemistry*, 1995, **38**, 2714-2727.
30. H. Tohma and Y. Kita, *Advanced Synthesis & Catalysis*, 2004, **346**, 111-124.
31. M. Frigerio, M. Santagostino and S. Sputore, *Journal of Organic Chemistry*, 1999, **64**, 4537-4538.
32. R. E. Ireland and L. Liu, *Journal of Organic Chemistry*, 1993, **58**, 2899-2899.
33. K. E. Pfitzner and J. G. Moffatt, *Journal of the American Chemical Society*, 1963, **85**, 3027-3028.
34. J. R. Parikh and W. v. E. Doering, *Journal of the American Chemical Society*, 1967, **89**, 5505-5507.
35. V. I. Stenberg and R. J. Perkins, *Journal of Organic Chemistry*, 1963, **28**, 323-324.
36. E. J. Corey and A. Palani, *Tetrahedron Letters*, 1995, **36**, 3485-3488.
37. D. Ma and W. Zhu, *Synlett*, 2006, **2006**, 1181-1184.
38. F. A. Luzzio, D. Y. Duveau and W. D. Figg, *Heterocycles*, 2006, **70**, 321-334.
39. B. S. Dyson, J. W. Burton, T.-i. Sohn, B. Kim, H. Bae and D. Kim, *Journal of the American Chemical Society*, 2012, **134**, 11781-11790.
40. J. Tsuji, *Synthesis*, 1984, **1984**, 369-384.
41. Y.-F. Wang, Y.-R. Gao, S. Mao, Y.-L. Zhang, D.-D. Guo, Z.-L. Yan, S.-H. Guo and Y.-Q. Wang, *Organic Letters*, 2014, **16**, 1610-1613.
42. C. Ballatore, D. M. Huryn and A. B. Smith, *ChemMedChem*, 2013, **8**, 385-395.
43. Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame and J. Taunton, *Journal of the American Chemical Society*, 2017, **139**, 680-685.
44. A. Wissner, M. Brawner Floyd, S. K. Rabindran, R. Nilakantan, L. M. Greenberger, R. Shen, Y.-F. Wang and H.-R. Tsou, *Bioorganic & Medicinal Chemistry Letters*, 2002, **12**, 2893-2897.
45. V. Chandregowda, G. Venkateswara Rao and G. Chandrasekara Reddy, *Heterocycles*, 2007, **71**, 39-48.
46. A. V. Statsuk, D. J. Maly, M. A. Seeliger, M. A. Fabian, W. H. Biggs, D. J. Lockhart, P. P. Zarrinkar, J. Kuriyan and K. M. Shokat, *Journal of the American Chemical Society*, 2008, **130**, 17568-17574.
47. R. R. Yadav, S. K. Guru, P. Joshi, G. Mahajan, M. J. Minto, V. Kumar, S. S. Bharate, D. M. Mondhe, R. A. Vishwakarma, S. Bhushan and S. B. Bharate, *European Journal of Medicinal Chemistry*, 2016, **122**, 731-743.
48. M. Zhao, H. Ning, M. Feng, S. Li, J. Chang and C. Qi, *Molecules*, 2014, **19**, 5508-5521.
49. F. Sanchez-Viesca and M. Berros, *Heterocycles*, 2002, **57**, 1869-1879.
50. Z. Fu, Z. Li, Y. Song, R. Yang, Y. Liu and H. Cai, *Journal of Organic Chemistry*, 2016, **81**, 2794-2803.
51. C. V. Wilson, *Journal of the American Chemical Society*, 1948, **70**, 1901-1903.

Chapter 5 – 1,4-Dicarbonyl Electrophiles Targeting the EGFR Catalytic Lysine Residue Lys745

52. Y. Loidreau and T. Besson, *Tetrahedron*, 2011, **67**, 4852-4857.
53. S. Bräse, H. Wertal, D. Frank, D. Vidović and A. d. Meijere, *European Journal of Organic Chemistry*, 2005, **2005**, 4167-4178.
54. M. Mori, Y. Washioka, T. Urayama, K. Yoshiura, K. Chiba and Y. Ban, *Journal of Organic Chemistry*, 1983, **48**, 4058-4067.
55. L. Gilfillan, R. Artschwager, A. H. Harkiss, R. M. J. Liskamp and A. Sutherland, *Organic & Biomolecular Chemistry*, 2015, **13**, 4514-4523.
56. F. Weber, S. Brune, K. Korpis, P. J. Bednarski, E. Laurini, V. Dal Col, S. Pricl, D. Schepmann and B. Wünsch, *Journal of Medicinal Chemistry*, 2014, **57**, 2884-2894.
57. J. E. Sieser, R. A. Singer, J. D. McKinley, D. E. Bourassa, J. J. Teixeira and J. Long, *Organic Process Research & Development*, 2011, **15**, 1328-1335.
58. R. M. Williams and M. N. Im, *Journal of the American Chemical Society*, 1991, **113**, 9276-9286.
59. M. Maiti, M. Maiti, J. Rozenski, S. De Jonghe and P. Herdewijn, *Organic & Biomolecular Chemistry*, 2015, **13**, 5158-5174.
60. J. C. Lukesh, M. J. Palte and R. T. Raines, *Journal of the American Chemical Society*, 2012, **134**, 4057-4059.
61. M. Jörres, I. Schiffers, I. Atodiresei and C. Bolm, *Organic Letters*, 2012, **14**, 4518-4521.
62. X.-H. Tan, B. Shen, W. Deng, H. Zhao, L. Liu and Q.-X. Guo, *Organic Letters*, 2003, **5**, 1833-1835.
63. A. Kolleth, M. Cattoen, S. Arseniyadis and J. Cossy, *Chemical Communications*, 2013, **49**, 9338-9340.
64. S. C. Ensign, E. P. Vanable, G. D. Kortman, L. J. Weir and K. L. Hull, *Journal of the American Chemical Society*, 2015, **137**, 13748-13751.

Chapter 6

Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Abstract

The research efforts described in this chapter continue with the development of inhibitors potentially capable of irreversibly targeting the EGFR catalytic Lys745. Acting as a catalyst for this work, the first reported crystal structure of EGFR, in which the catalytic Lys745 was covalently modified, will be reviewed. Inspired by this, we chose to adapt our therapeutic targeting strategy to include the successful sulfonyl fluoride electrophile, as well as other lysine targeting electrophiles including ethenesulfonamide, nitriles, aliphatic chain sulfonyl fluorides and α - and β -haloketones. Perusal of the literature and use of molecular modelling led selection of an osimertinib-derived scaffold as prime candidate for our new heterocyclic driving group. The synthesis of this heterocyclic driving group, proving arduous and requiring exploration of several synthetic strategies and pathways, will be described in full. Ultimately, we were able to develop an optimised route with high overall yield, with five intermediate compounds submitted for evaluation. Installation of the chosen electrophiles proved challenging, with numerous contributing factors identified and rationalised. Overcoming these difficulties led to the successful synthesis of four final compounds, whose inhibitory properties were measured in subsequent biochemical and cellular evaluation. While ineffective in covalently modifying the catalytic lysine residue, the assessment of several inhibitors demonstrated intriguing results which will be disclosed. Finally, improvements of synthetic methodology and expansion toward new electrophiles and driving group scaffolds are considered.

6.1 Introduction

As stated in the concluding paragraph of the previous chapter, the following body of research was instigated to a large extent by work undertaken by Taunton and co-workers.¹ As protein kinases comprise one of the largest families of structurally related enzymes, a major aim in the development of their inhibitors lies in selectively targeting a singular desired kinase, whilst avoiding the myriad of potential off-target kinases. However, the capacity to accurately quantify inhibitor interactions with multiple off-target kinases in live cells remains problematic. The Taunton research group was able to design and successfully synthesise several chemical probes that covalently labelled a broad spectrum of the intracellular kinome with high efficiency. These clickable piperazine-linked probes, illustrated in **Figure 6.1a** as **1** – **3**, made use of an electrophilic sulfonyl fluoride moiety attached to a pyrimidine 3-aminopyrazole driving group.¹ This scaffold can form three hydrogen bond interactions with the hinge region of the kinase, and as such, has been shown to inhibit an extensive number of kinases, including those with bulky gatekeeper residues such as the EGFR-T790M mutant.²

Remarkably, the optimized probe **2** managed to capture up to 133 endogenous kinases from a single cell line, with the previous best pan-kinase probe, a photoreactive diazirine, only able to detect up to 22 intracellular kinases.³ Furthermore, probe **2** was able to compete effectively with high concentrations of intracellular ATP in these demanding chemoproteomic applications, highlighting its utility. Use of protein crystallography and mass spectrometry experiments confirmed that covalent modification occurred chemoselectively with the conserved catalytic lysine residue within the ATP binding site. Of the labelled 133 kinases, the first covalently modified EGFR wild-type catalytic Lys745 residue co-crystal structure was solved and is shown on the following page in **Figure 6.1b**.

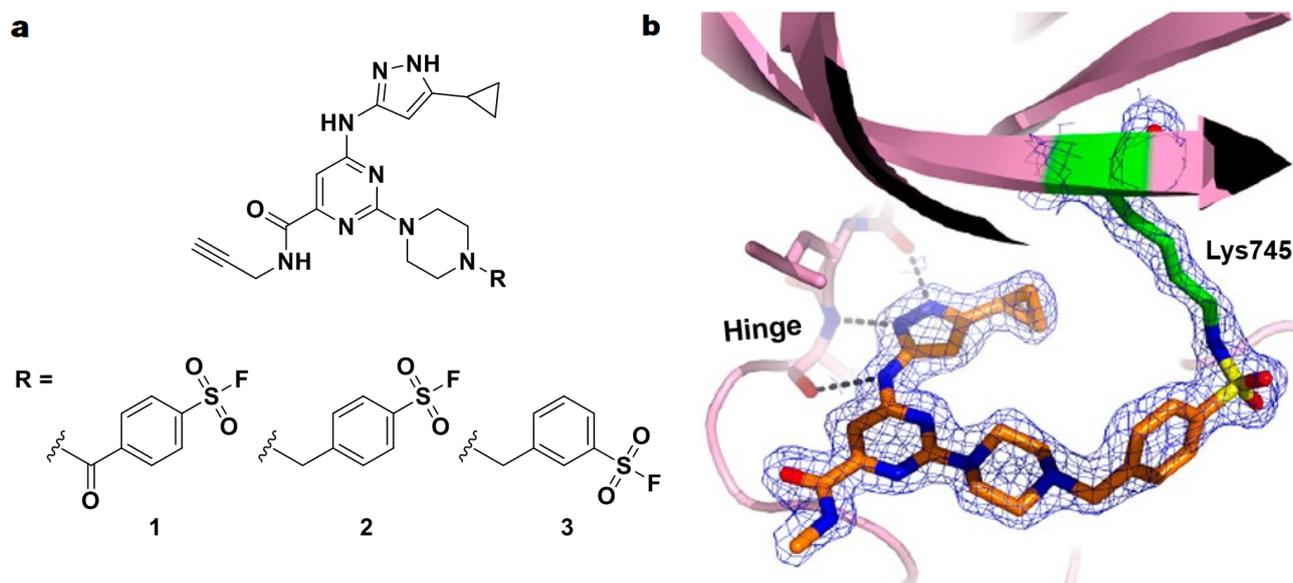


Figure 6.1: a) Pyrimidine 3-aminopyrazole based compounds **1** - **3** with piperazine linked sulfonyl fluorides acting as lysine-targeting chemical probes, b) Co-crystal structure of probe **2** (orange) covalently bound to the catalytic Lys745 (green) through a sulfonyl fluoride electrophile (PDB: 5U8L). Figure reproduced from b) Zhao et al.¹

The announcement that the EGFR catalytic lysine could indeed be covalently modified was a welcome and encouraging revelation. However, this compelled us to re-evaluate our previously implemented strategy of using 1,4-dicarbonyl warheads to potentially undergo Paal-Knorr pyrrole formation with the catalytic lysine. Seeking to emulate the success of Taunton's research group in targeting and covalently modifying the catalytic Lys745 residue, we envisaged replacement of the 1,4-dicarbonyl electrophiles with the implemented reactive sulfonyl fluoride warheads. Furthermore, we considered the selection of a quinazoline-based driving group inadequate to successfully compete with and overcome the EGFR-T790M and -C797S mutant variants, which possess a restored and increased binding affinity for endogenous ATP.⁴ While probe **2** displayed a mediocre wild-type EGFR inhibition of 36% at 1 μM , we believed that the use of an optimised and selective EGFR heterocyclic driving group, in combination with electrophiles such as the sulfonyl fluoride employed in probe **2**, had the potential to produce therapeutically relevant lysine-targeting irreversible inhibitors. Perusal and a structural analysis search of the literature led us to several prospective electrophiles and heterocyclic scaffold candidates, which will be discussed in the following section.

6.2 Electrophile and Heterocyclic Scaffold Selection

6.2.1 Additional Lysine-trapping Electrophiles

6.2.1.1 Sulfonyl Fluorides

Sulfonyl fluorides are hydrolysis-resistant electrophiles that have found substantial utility as reactive probes in chemical biology and molecular pharmacology.⁵ They are considered privileged warheads, owing to their inherent balance of favourable biocompatibility and protein reactivity. As such, they are known to modify not only serine (particularly as protease inhibitors) but also lysine, threonine, histidine and cysteine residues in a context specific manner.⁶

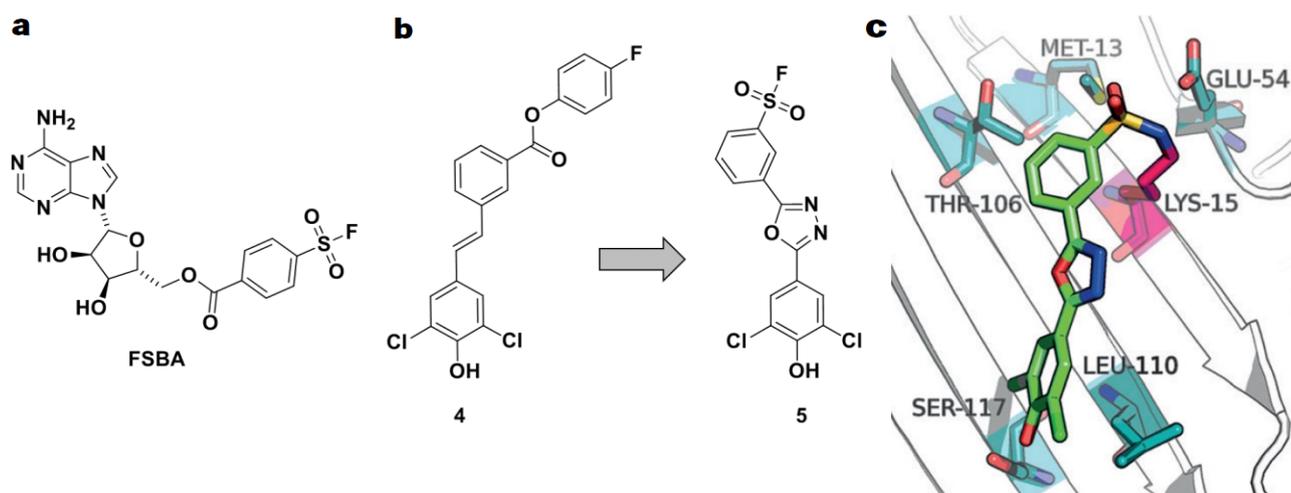


Figure 6.2: a) Chemical probe **FSBA**, with sulfonyl fluoride electrophile, b) TTR stabilizer **4** and its aromatic sulfonyl fluoride derived analogue **5**, c) Crystal structure of **5** (green) covalently bound to Lys15 (pink) through the aromatic sulfonyl fluoride (PDB: 4FI7). Figure reproduced from c) Pettinger, Jones and Cheeseman.⁷

Despite their desirable traits, sulfonyl fluorides have seldom been employed in the structure-based design and development of lysine-targeted covalent inhibitors. Originally developed 40 years ago, Colman et al. designed and pioneered the use of the adenosine derivative 5'-fluorosulfonylbenzoyl 5'-adenosine (**FSBA**), shown in **Figure 6.2a**, to explore nucleophilic residues in the glutamate dehydrogenase binding site.⁸ The ATP analogue contains an aromatic sulfonyl fluoride functionality that replaces the phosphoryl group of ATP and serves as a hard electrophile able to readily react with a pK_a -perturbed lysine residue, allowing for covalent bond formation.⁷ FSBA has become frequently used as a covalent inhibitor and probe of ATP-binding enzymes, serving as inspiration for the 1,4-dialdehyde probe discussed in **Section 5.2.2** and the chemical probes **1 – 3**.

Outside of kinases, sulfonyl fluoride inhibitors of transthyretin (TTR) have been developed to react covalently with the pK_a perturbed Lys15 in the thyroxine binding site, thereby stabilizing the enzyme and preventing fibrillation.⁹ Initially, Kelly et al. adapted a reversible stabilizer of TTR to afford compound **4**, illustrated in **Figure 6.2b**, which was able to covalently modify the non-catalytic Lys15 via the activated ester.¹⁰ Further development led to the sulfonyl fluoride-containing compound **5**, which displayed superior reactivity towards covalent bond formation. This was hypothesised to proceed via a stabilised Glu54 mediated transition-state, allowing for effective covalent inhibition of Lys15 as shown in **Figure 6.2c**.⁹ These examples highlight the potential utility of sulfonyl fluorides in covalent enzyme inhibition.

6.2.1.2 Michael Acceptors

Michael acceptors have traditionally been tailored to target a soft nucleophilic cysteine residue within the ATP binding site. For covalent lysine modification, harder electrophiles are expected to react more readily with the hard ϵ -amine of lysine.¹¹ To evaluate the inherent reactivity profile of lysine, Dahal and co-workers determined the rate reaction of *N*- α -acetyl-lysine as nucleophile with several electrophilic moieties typically used as irreversible inhibitor warheads, the results of which is tabulated in **Figure 6.3a**.¹² To best mimic the hypothetical perturbed pK_a of a lysine within the protein microenvironment, the assessment was carried out at pH 10.2 and compared with the reactivity of the soft thiol nucleophile glutathione (GSH).

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

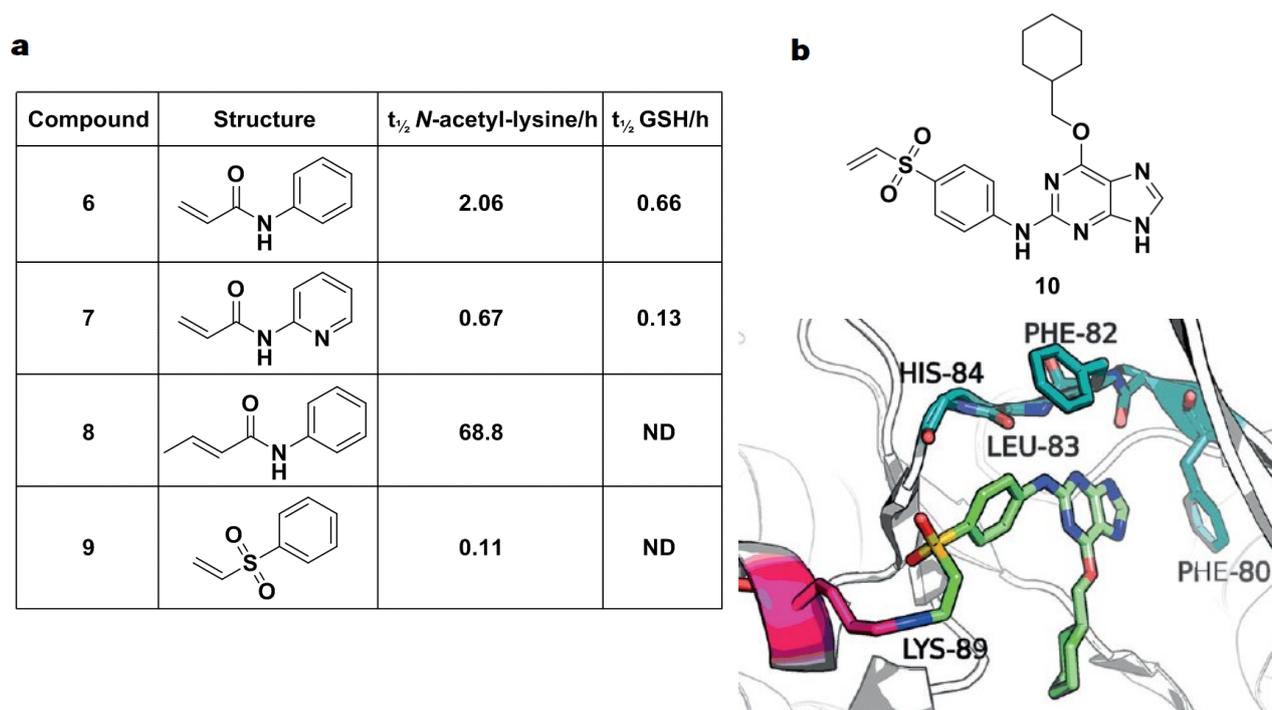


Figure 6.3: a) Tabulated reactivities of hard and soft electrophiles with *N*- α -acetyl-lysine and GSH, b) Structure of compound **10**, which was able to covalently modify the CDK2 catalytic Lys89, as shown in the crystal structure (PDB: 5CYI). Figure reproduced from b) Pettinger, Jones and Cheeseman.⁷

The lysine nucleophile displayed good reactivity with the conventional acrylamide **6**, with reactivity increasing relative to additional electron-withdrawing effects, as seen with pyridine **7**. As expected, this trend was more pronounced for the soft GSH. Introduction of a methyl group at the vinyl terminus (**8**) drastically slowed the reaction with the lysine nucleophile due to steric hindrance. Interestingly, the vinyl sulfone **9** exhibited the greatest reactivity for the *N*- α -acetyl-lysine and apparent selectivity over GSH. This was hypothesised to stem from the net negative charge of the sulfone, leading to a more polarizable vinyl moiety and significant decrease in softness.¹²

Based upon previous work which identified a series of potent reversible inhibitors of CDK2, Anscombe et al. identified a proximal interaction between the inhibitor's sulfonamide moiety and Lys89. Consequently, the group synthesised the vinyl sulfone analogue **10**, shown in **Figure 6.3b**, which was confirmed to bind covalently to Lys89 through protein crystallography and single-point mutation studies.¹³ This was the first example of an irreversible CDK2 inhibitor and successful application of this warhead in kinases, with compound **10** exhibiting superior cellular activity over its reversible counterpart. We found this particularly intriguing, as this highlighted the vinyl sulfone as a potential electrophile in targeting the EGFR catalytic Lys745.

6.2.1.3 Nitriles and α -Haloketones

In recent decades, a continual stream of nitrile-containing pharmaceuticals has resulted in over 30 agents being prescribed for a diverse variety of diseases and more than 20 additional candidates in clinical trials.¹⁴ This prevalence attests to the functional groups advantages and biocompatibility in drug design. These moieties are unconventional by nature due to their short polarised triple bond, which has been described as being 8 times smaller than that of a methyl group, resulting in a miniscule steric demand along its axis.¹⁴

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

They are also fairly robust and not readily metabolized, mostly passing through the body unchanged.¹⁵ While nitriles are not principally electrophilic to free nucleophiles, activation by adjacent structural components, such as electron-withdrawing groups, has been shown to facilitate covalent modification.¹⁶ Furthermore, nitriles often act as hydrogen bond acceptors with various amino acid residue such as lysine, serine or arginine.¹⁷ For these reasons, we wished to incorporate a nitrile functionality within our library of compounds to either form favourable reversible interactions or potentially undergo covalent modification of the EGFR catalytic Lys745 residue.

Lastly, we propose the utilisation of α -haloketone electrophiles, which have been shown to undergo nucleophilic attack by oxygen, nitrogen and sulphur nucleophiles.¹⁸ Owing to their high reactivity, haloketones take part in several reaction types and serve as versatile synthons in the synthesis of various classes of heterocycles.¹⁹ In reactions with nucleophiles, two electrophilic sites are available - the carbonyl functionality itself or the carbon carrying the halogen atom. The inherent reactivity of the α -haloketone stems from the inductive effect of the carbonyl group. This enhances the polarity of the carbon-halogen bond, facilitating increased electron deficiency and electro positivity of the α -carbon. Furthermore, this feature is augmented with the more polar carbon-halogen bond, allowing for faster reaction with nucleophiles.¹⁹ Consequently, we imagined the use of this class of electrophile to potentially label the catalytic lysine residue of EGFR.

6.2.2 Heterocyclic Scaffolds Exploiting Catalytic Lysine Interactions

In a recent study, Katayama and co-workers found the dual ALK/EGFR inhibitor brigatinib to be an effective inhibitor of EGFR-C797S triple mutant-harbouring cells, both *in vitro* and *in vivo*.²⁰ While the potency of the drug as monotherapy was found to be unsatisfactory, its combination with anti-EGFR antibodies dramatically increased the efficacy by inducing surface degradation of the receptor and halting EGFR overexpression. This in turn led to a decrease in tumour size of C797S mutation-bearing cells and a prolonged survival period for xenograft mice. Molecular modelling simulations of brigatinib within the ATP binding site, revealed a compatible binding affinity and several key components which have potential for further development. The group believed this could afford more potent analogues capable of combating the therapeutically relevant EGFR triple mutant.

Building on these efforts, a publication by Jang et al. rationally designed and synthesised a host of dual inhibitors of ALK and EGFR.²¹ This was carried out by combining the critical functional elements of the ALK inhibitors, ceritinib and brigatinib, with features of the EGFR inhibitors osimertinib and WZ4002, shown in **Figure 6.4a**. Whilst retaining the 2,4-diarylaminopyrimidine scaffold, the use and interchange of numerous solubilising and other functional groups such as isopropylsulfone, dimethylsulfonamide and dimethylphosphine oxide were investigated. This resulted in the identification of **11** as the most potent dual ALK/EGFR inhibitor, which exhibited low nanomolar IC₅₀ values biochemically and maintained activity in cellular growth inhibition studies. As such, compound **11** has been identified as a lead candidate, with further optimization to improve metabolic stability and kinase selectivity currently underway.

One of the most striking features of **11**, when modelled and visualised within the ATP binding pocket, is the proximity and interactions of the isopropylsulfone moiety with the catalytic Lys745 residue, as shown in **Figure 6.4b**.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

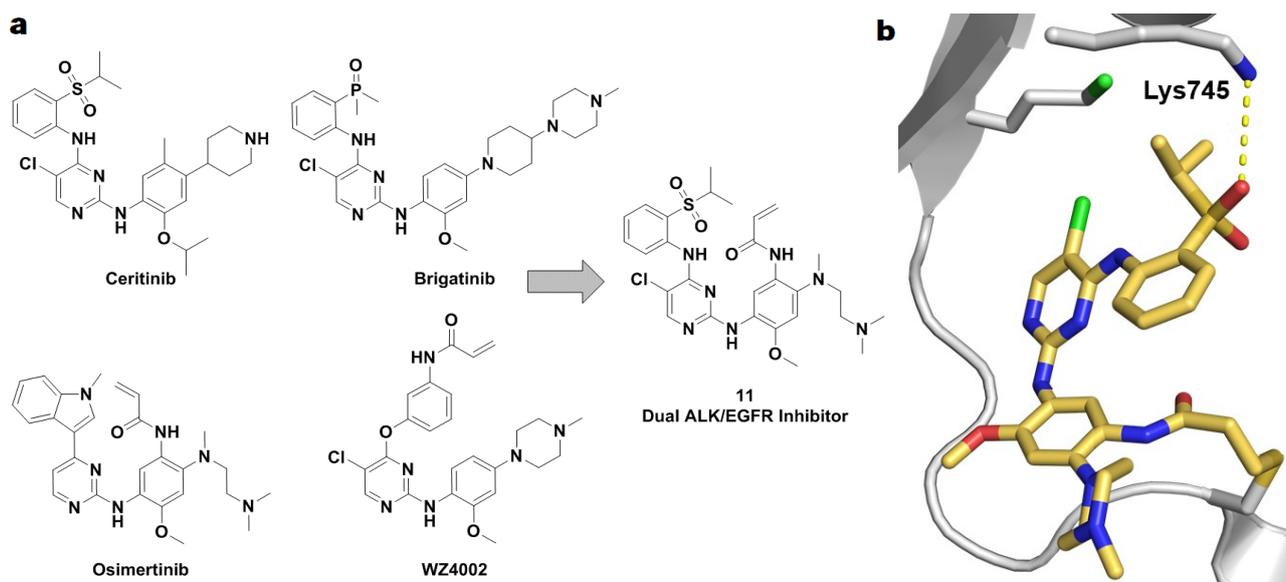


Figure 6.4: a) Structures of the ALK inhibitors ceritinib and brigatinib, EGFR inhibitors osimertinib and WZ4002 and the most efficacious dual ALK/EGFR inhibitor compound **11**, b) Compound **11** docked into the EGFR-T790M double mutant, illustrating hydrogen bond interactions between the isopropylsulfone group and the catalytic Lys745 residue (PDB: 5GTZ).

Similar to the dimethylphosphine oxide in brigatinib, the carbonyl groups are able to act as hydrogen bond acceptors and form favourable reversible interactions with Lys745. This was particularly importance to us, as substitution of the isopropylsulfone group with an appropriate electrophile, such as a sulfonyl fluoride, could potentially induce covalent bond formation with this residue. We therefore envisioned use of the 2,4-diarylaminopyrimidine scaffold in the design of our lysine targeting inhibitors.

Osimertinib displays a 200-fold selectivity for the L858R/T790M double mutant over wild-type EGFR.²² In an attempt to understand the molecular basis of this selectivity, many groups have sought to produce crystal structures of the inhibitor within the wild-type, T790M and L858R/T790M variants of the enzyme. In 2015, a crystal structure of osimertinib bound to the wild-type EGFR enzyme was solved and is shown in **Figure 6.5a**.²² While the binding mode of the blockbuster drug was revealed for the wild-type, the researchers could unfortunately not complex the inhibitor with the single activating and double mutant variants. This remains the case to date, with no crystal structure of osimertinib bound to the mutated enzymes found in the literature or uploaded to the PDB.

Without a crystal complex of osimertinib bound to the EGFR-T790M kinase domain, Zhang et al. employed molecular modelling studies to determine a predicted binding mode of the FDA approved inhibitor.²³ From the 2D representation shown in **Figure 6.5b**, the pyrimidine core was found to act as a hydrogen bond donor and acceptor, thereby interacting with Met793 (red) of the hinge region and effectively evading steric interactions with the mutated Met790 residue (red). As expected, the acrylamide electrophile was found to covalently bind to Cys797 (green), whilst the dimethylamine moiety was appropriately protruding into the solvent channel (blue). Interestingly, the group observed the presence of a hydrophobic pocket (yellow) between the indole functionality and the specificity pocket of the kinase. One of these neighbouring residues was the catalytic Lys745 (purple), forming the well-known salt bridge with Glu762 (orange). This piqued our curiosity, leading us to run our own modelling simulations of osimertinib within the ATP binding site. From this, we calculated the distance between the N-methyl carbon of the indole and the charged amine of Lys745 to be 5.39 Å.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

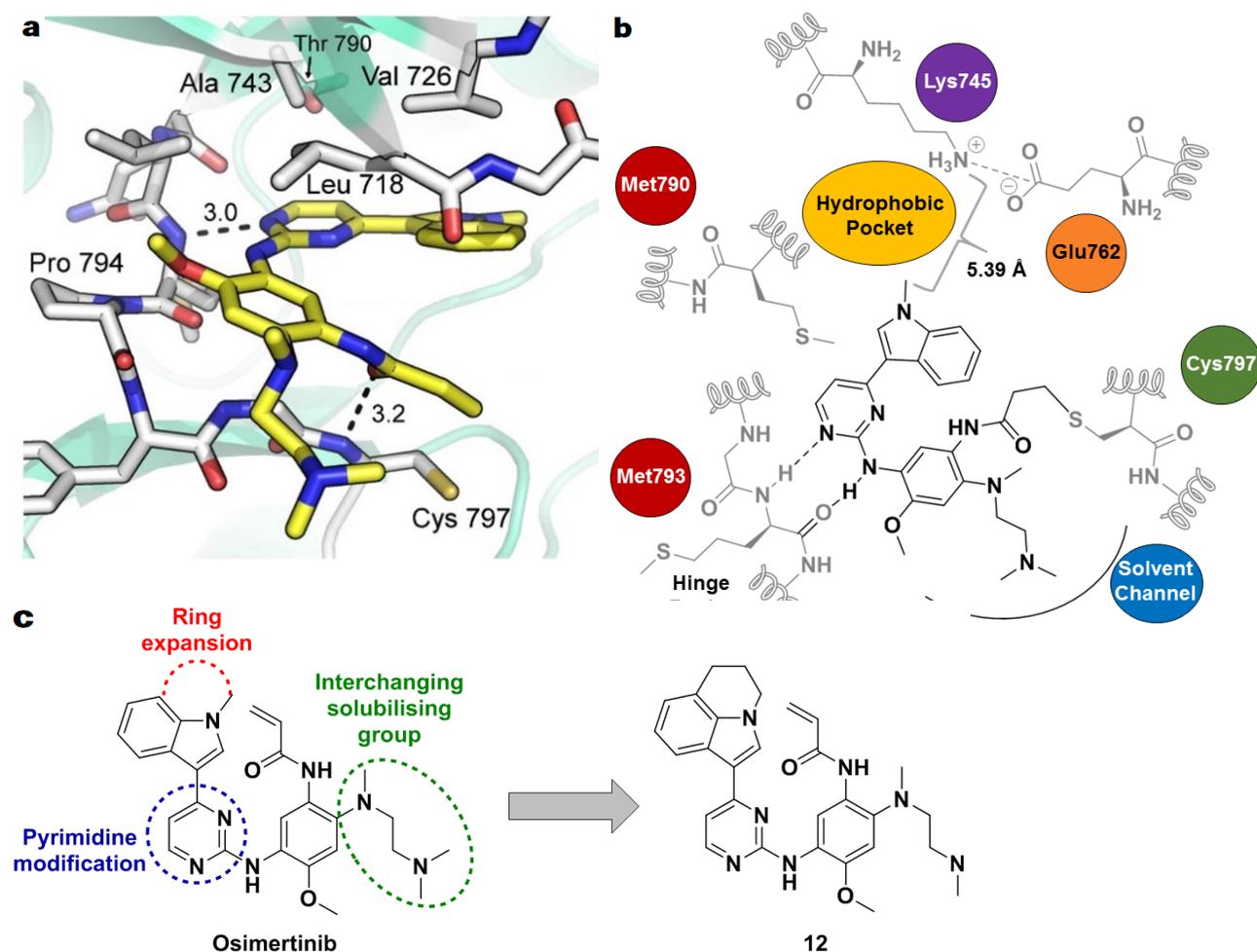


Figure 6.5: a) Crystal structure of osimertinib (yellow) bound to the wild-type EGFR domain (PDB: 4ZAU), b) 2D representation of osimertinib modelled into the EGFR L858R/T790M double mutant, with selected van der Waals and covalent interactions with specific amino acids highlighted in colour (PDB: 3IKA), c) SAR investigation of osimertinib affording the most potent compound **12**. Figure reproduced from a) Yosaatmadja et al.²²

Based on the modelled binding mode of osimertinib with the T790M mutant, the group undertook a SAR investigation to clarify the inherent selectivity of the inhibitor. This included modification of the substitution position of the pyrimidine ring, interchange of the amine solubilising group, and importantly, expansion of the indole ring into the exposed hydrophobic pocket as shown in **Figure 6.5c**. As expected, osimertinib was found to be optimised in all parameters except one – the expansion of the indole ring. This led the group to the 5,6-dihydro-4*H*-pyrrolo[3,2,1-*ij*]quinoline compound **12**, which retained excellent selectivity and sub-nanomolar IC₅₀ activity values for the single L858R and double L858R/T790M mutants. Furthermore, compound **12** displayed enhanced microsomal stability and pharmacokinetic properties, with a lower binding affinity to the human *ether-à-go-go*-related gene (hERG) ion channel than osimertinib, resulting in lower cardiotoxicity.²³ On the basis of this experimental investigation, we envisaged the potential for covalent modification of the adjacent catalytic Lys745. We believed this could be accomplished through the installation of a suitable electrophile at the indole nitrogen of the osimertinib scaffold. With this intention in mind, we designed a library of compounds to be synthesised which is described in the following section.

6.3 Objectives and Target Synthetic Library

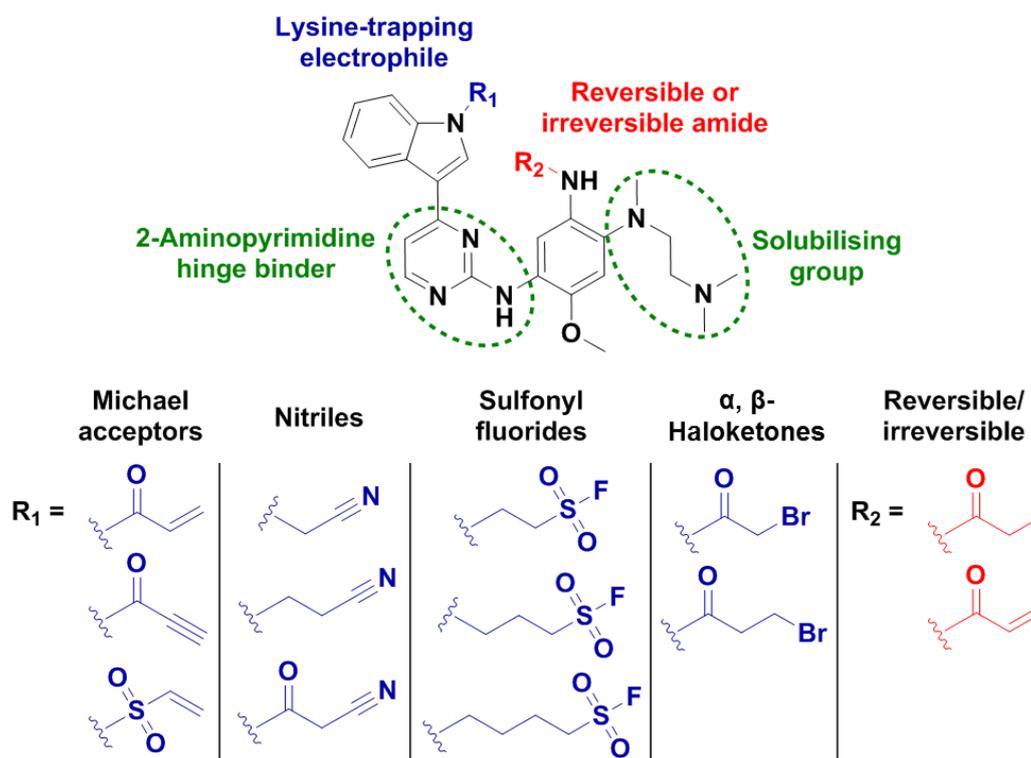


Figure 6.6: Target synthetic library for osimertinib derived catalytic lysine-targeting inhibitors.

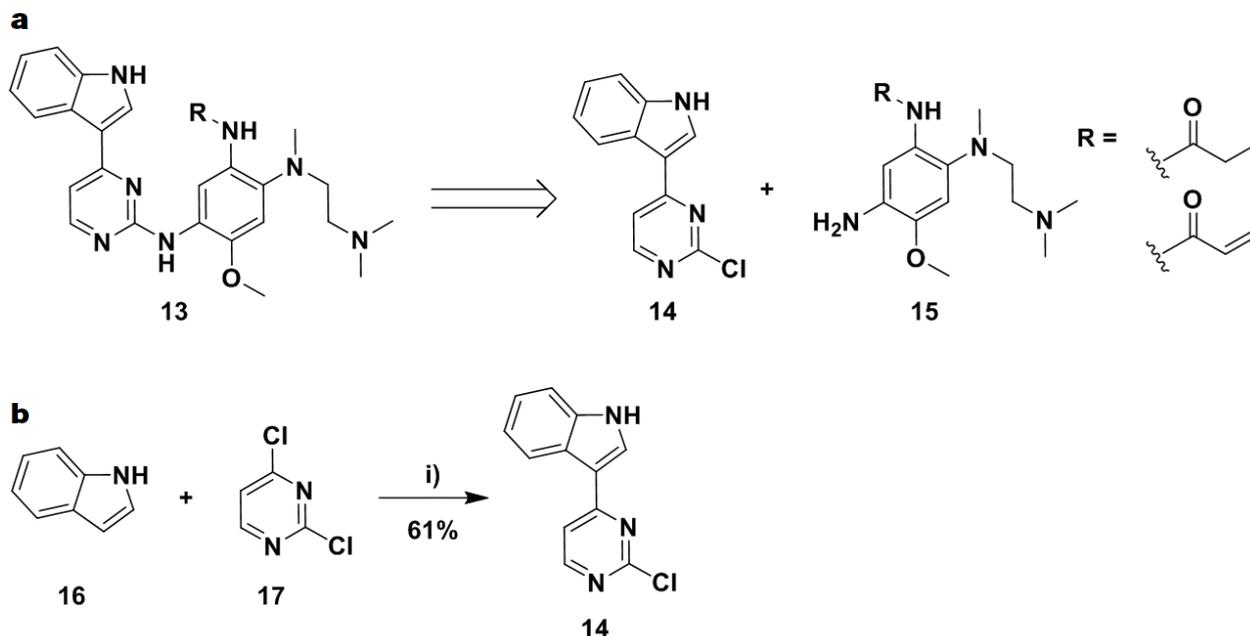
On the grounds of the previous section's discourse, we opted to employ the osimertinib heterocyclic driving group in the synthesis of our lysine-targeting library of compounds. The design elements of this library are illustrated above in **Figure 6.6**. Utilisation of this scaffold was considered desirable for several reasons, the first being inclusion of the 2-aminopyrimidine motif (green) as hinge binder, which has been shown to be effective in inhibitors of both ALK and EGFR that were able to overcome resistance mutations.²¹ Furthermore, as established by Zhang et al. in their SAR studies and by Jang and co-workers in the synthesis of the dual ALK/EGFR inhibitor **11**, the solubilising group (green) and other constituents of osimertinib were found to be optimised, leading us to use an otherwise unaltered structure.^{21, 23}

For the lysine-targeting electrophiles, we envisaged derivatisation at the indole nitrogen (blue). On the basis of our molecular modelling, we proposed that incorporation at this position could theoretically place the warhead in a favourable orientation and proximal to the catalytic Lys745, increasing the likelihood of covalent bond formation. These electrophilic functionalities, shown in blue in **Figure 6.6**, would include the previously highlighted Michael acceptors, nitriles, sulfonyl fluorides and α-, β-haloketones.

In conjunction with this, we wished to synthesise the reversible propionamide and irreversible acrylamide derivatives of osimertinib, highlighted in red in **Figure 6.6**. Understandably, this could provide valuable insight and a comparative platform in the case of covalent bond formation at the catalytic lysine residue. We therefore set out in the pursuit of these target compounds, in the hope of covalently modifying the EGFR catalytic Lys745 and thereby improving the efficacy and selectivity against the C797S mutant.

6.4 Library Synthesis

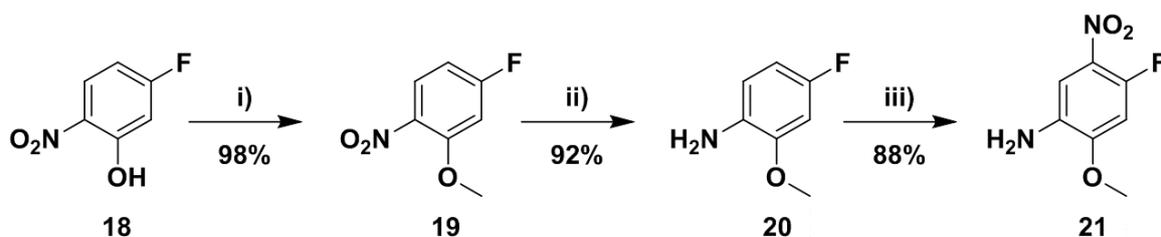
6.4.1 Osimertinib-derived Scaffold Synthesis



Scheme 6.1: a) Retrosynthesis of scaffold **13** and b) Synthesis of heterocyclic fragment **14**. Reagents and conditions: i) indole (1.5 equiv.), 2,4-dichloropyrimidine (1 equiv.), EtMgBr (generated *in situ*) (1.5 equiv.), THF, $-10\text{ }^\circ\text{C} - 80\text{ }^\circ\text{C}$, 5 h.

For the synthesis of our osimertinib based lysine-targeting inhibitors, we envisaged heterocyclic scaffold **13**, seen in **Scheme 6.1** above, to serve as a platform for further electrophile derivatisation. Preliminary retrosynthetic analysis of **13** led us to a disconnection at the pyrimidine aniline bond and the two heterocyclic fragments **14** and **15**. Successful synthesis of the indole-pyrimidine **14** and the acryl- and propionamide variant of fragment **15** would set up for a subsequent $\text{S}_{\text{N}}\text{AR}$ reaction between these two fragments to afford the desired **13**. We believed that execution of this strategy would accommodate a more divergent synthesis, allowing for the combination of **14** with other aniline fragments such as **15** in future synthetic endeavours. We therefore set about designing and implementing a synthetic pathway to these two fragments.

For coupling of the commercially available indole (**16**) and 2,4-dichloropyrimidine (**17**) at the 4-position, we employed the use of a Grignard reagent to act as base. Deprotonation of the *N*-indole proton would negate its participation as a nucleophile in the reaction while simultaneously activating the 3-position of the indole. The Grignard reagent used in this reaction was formed *in situ* with an excess of magnesium turnings and iodine as catalyst, with the appropriate alkyl halide. Initial attempts using methylmagnesium bromide in DCE furnished inadequate and unreliable yields, necessitating optimization of the reaction parameters. Optimization included use of various solvents, temperatures and more substituted alkyl halides in Grignard reagent formation. By using a 1.5 equivalent excess of ethylmagnesium bromide and indole while heating under reflux in THF, we were able to repeatedly reproduce a yield of roughly 60%, a slight improvement on that provided in the literature.²⁴ Furthermore, these reactions could be scaled up to 8 grams, thereby affording us with a large amount of the heterocyclic building block **14** to be used for future coupling reactions.



Scheme 6.2: Three step synthesis of intermediate fragment **21**. Reagents and conditions: i) methyl iodide (1.1 equiv.), K_2CO_3 (1.2 equiv.), acetone, $0\text{ }^\circ\text{C} - 60\text{ }^\circ\text{C}$, 12 h; ii) H_2 , Pd/C (10 mol%), EtOH, rt, 24 h; iii) KNO_3 (1 equiv.), 85% H_2SO_4 (25 equiv.), $-10\text{ }^\circ\text{C} - \text{rt}$, 4 h.²⁴

The first steps toward the other heterocyclic building block proceeded smoothly and commenced with methylation of the commercially available phenol **18**, as seen above in **Scheme 6.2**. Use of methyl iodide and potassium carbonate as base in acetone provided a near quantitative yield of **19**, with no further purification required. This was followed by aromatic nitro reduction using the standard conditions of Pd/C with hydrogen atmosphere in ethanol, to provide compound **20** in excellent yield. Lastly, regioselective nitration was carried out in slightly diluted sulphuric acid by portion-wise addition of potassium nitrate, while maintaining a temperature below $-5\text{ }^\circ\text{C}$. This afforded the intermediate fragment **21** in good yield, with ^1H NMR spectrum shown in **Figure 6.7** below.

Furthest upfield we see a singlet at δ 3.90 ppm which correctly integrates for 3H and thus corresponds to the methoxy peak (blue). Whilst the combined activating/deactivating effects of the ring's substituents suggest nitration at the 2-position, we propose the steric interactions of the tri-substituted aromatic product formed by this, to outweigh these effects. Analysis of the obtained ^1H NMR spectrum lends credence to this, with the broad amine peak (purple) appearing at the unchanged (relative to the starting material) position of δ 5.22 ppm. Were the nitro functionality ortho to the aniline, we would expect a large downfield shift in this peak due to the electron-withdrawing nature of the nitro group. Furthermore, we would expect the appearance of doublet of doublets in the aromatic region were nitration to occur at either the 2- or 5-position, as each of the available protons would be able to couple to both a neighbouring fluorine and proton.

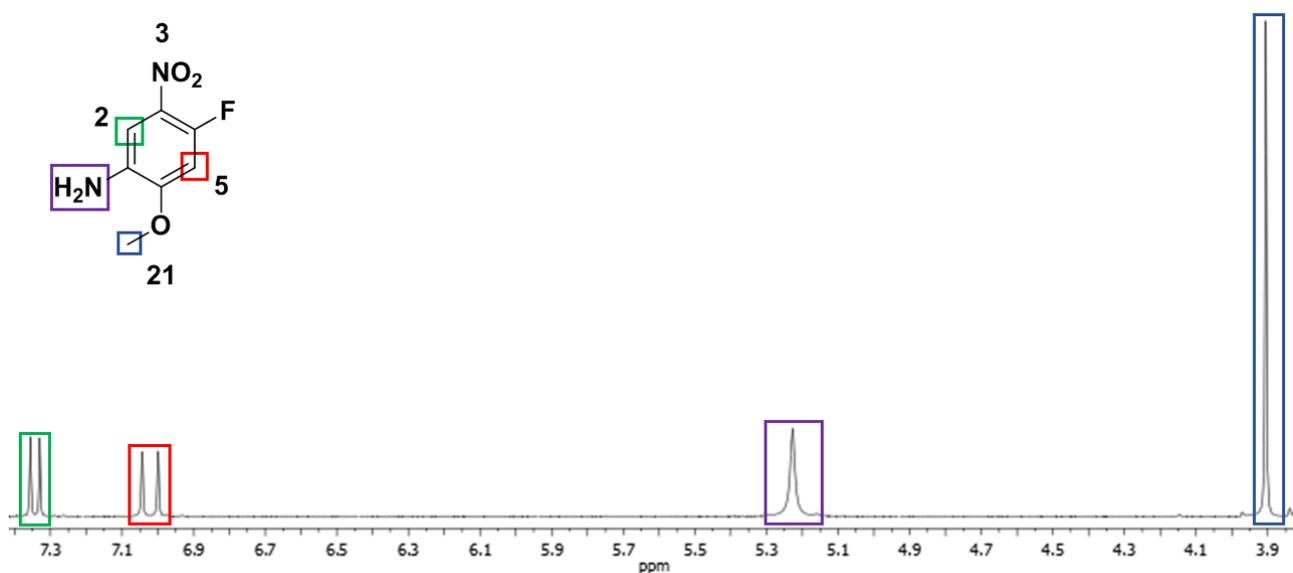


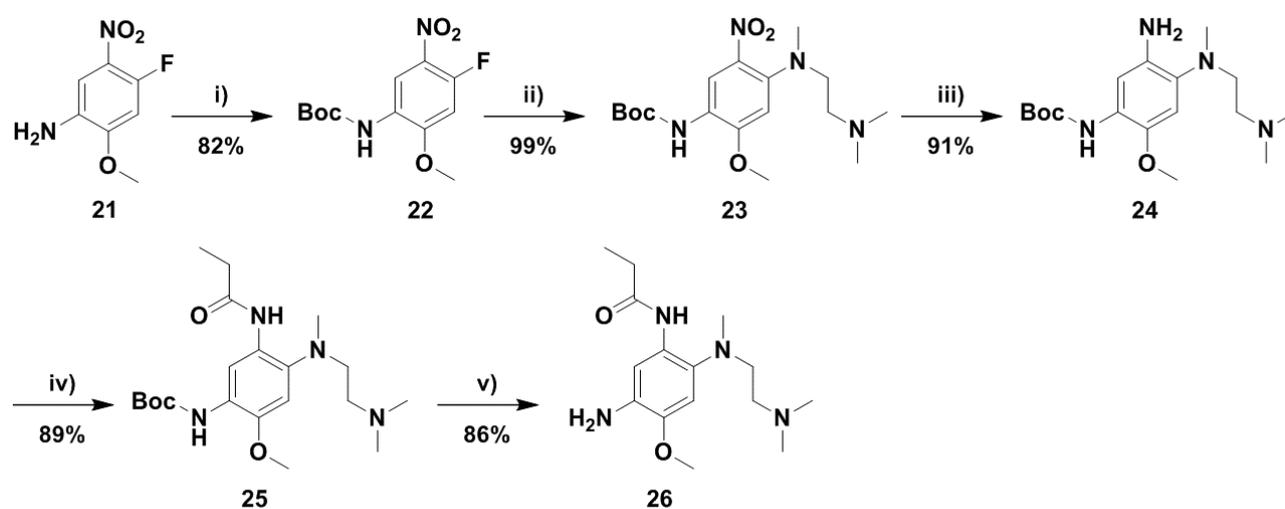
Figure 6.7: ^1H NMR spectrum of intermediate fragment **21**.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

However, the two aromatic peaks appeared as doublets, with the 2-position proton ($J = 7.8$ Hz, green) occurring more downfield in comparison to the 5-position proton ($J = 13.4$ Hz, red), as it is *ortho* to the electron withdrawing nitro moiety. Furthermore, the difference in J-coupling values are indicative of 2J (red) and 3J (green) H-F coupling to the fluorine atom. Lastly, the found HRMS $[M+H]^+$ ion of 187.0524 (calculated: 187.0519), 1H and ^{13}C NMR spectra all compared well with the reported literature values.²⁵

Continuing with the synthesis of the heterocyclic building block **26** in **Scheme 6.3**, we anticipated the required protection of the aniline group to avoid reaction centre competition in the forthcoming acylation. We therefore initiated the synthesis with a Boc-protection of compound **21**. Interestingly, our first attempts to introduce the Boc-protecting group using non-nucleophilic bases such as Et_3N and DIPEA resulted in the preferred formation of a couple of by-products. Due to the highly electronegative nature of the fluorine atom and its placement *ortho* to the electron-withdrawing nitro group, we postulated the occurrence of a rapid intermolecular S_NAr reaction between the nucleophilic aniline and another equivalent of starting material prior to protection. Formation of these by-products was witnessed even at low temperatures, leading us to avoid the use of bases which could expedite the incidence of the hypothesised S_NAr reaction. Following several optimization reactions, use of di-*tert*-butyl dicarbonate with a catalytic amount of DMAP in DCM afforded the Boc-protected analogue **22** in good yield, albeit over a long reaction time.

Introduction of the solubilising group in the subsequent S_NAr reaction between **22** and N^1,N^1,N^2 -trimethylethane-1,2-diamine proceeded flawlessly with DIPEA acting as base, furnishing a near quantitative yield of compound **23**. The reactivity exhibited in this conversion can be construed as validation for our previously outlined hypothesis concerning the difficulties with the Boc-protection. This allowed for aromatic nitro reduction of compound **23**, which was carried out using Pd/C under hydrogen atmosphere, with methanol discerned as the optimal solvent. Filtration through a plug of diatomaceous earth and purification of the crude product using flash column chromatography afforded compound **24** as a brown oil in excellent yield.



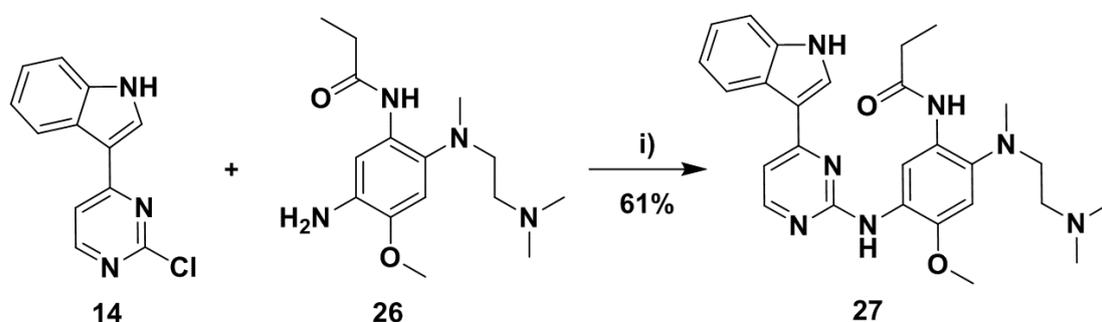
Scheme 6.3: Five step synthesis of heterocyclic fragment **26** to undergo coupling with **14**. Reagents and conditions: i) Boc_2O (1.05 equiv.), DMAP (cat.), DCM, rt, 24 h; ii) N^1,N^1,N^2 -trimethylethane-1,2-diamine (1.25 equiv.), DIPEA (1.05 equiv.), DMA, 60 °C, 4 h; iii) H_2 , Pd/C (10 mol%), MeOH, rt, 12 h; iv) propionyl chloride (1.05 equiv.), DIPEA (1.1 equiv.), DCM, -10 °C – rt, 2 h; v) DCM/TFA (3:1), rt, 6 h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Advancing with the synthetic route, we chose to progress with the propionamide version of the fragment for optimization purposes. Following the successful synthesis and future coupling reaction, we could return to build the acrylamide counterpart at a later stage. Accordingly, acylation of the aniline fragment **24** with propionyl chloride at low temperature in DCM furnished the propionamide **25** in a yield of 89%. To conclude the synthesis, deprotection of the Boc group was facilitated by a 3:1 mixture of DCM/trifluoroacetic acid, to afford compound **26** a beige solid. From compound **18**, the overall yield for the 8-step synthesis was calculated to be approximately 45%.

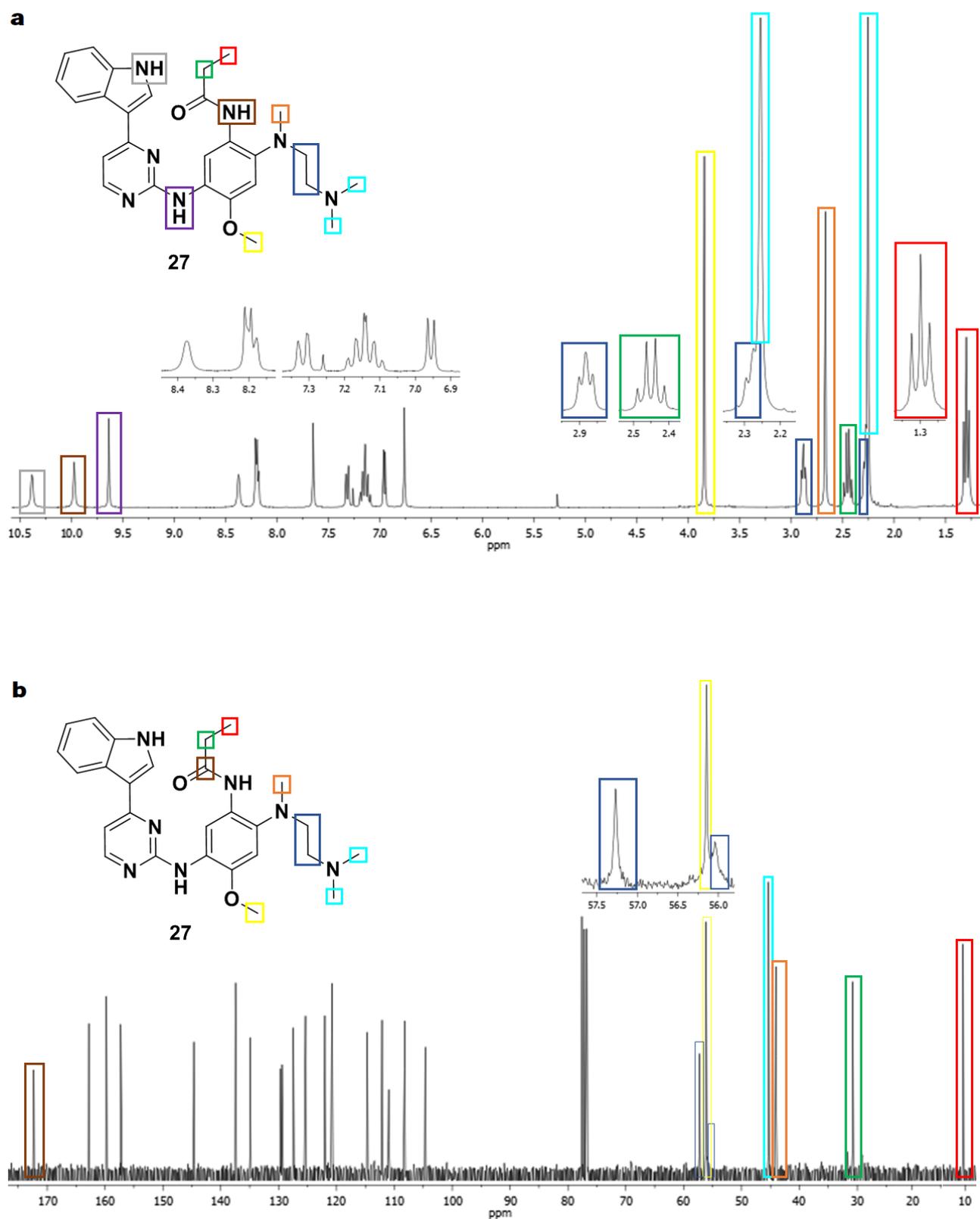
With the heterocyclic fragment **26** in hand we looked towards coupling with the previously synthesised indole-pyrimidine fragment **14**, illustrated in **Scheme 6.4** below. However, this reaction proved exceedingly problematic, requiring numerous attempts to optimize and attain the yield shown below. The reaction was found to proceed best under acidic conditions, requiring 1.2 equivalents of recrystallised *p*-toluenesulfonic acid monohydrate. Additionally, 2-pentanol was identified as the optimal solvent choice, determined from a pool of high boiling point solvents including, amongst others, NMP, ethylene glycol and DMSO. Application of these conditions afforded the osimertinib derived scaffold **27** in a maximum yield of 61%. Unfortunately, yields were found to be inconsistent and could only be reliably reproduced in small scale quantities of up to ~ 0.50 mmol. Furthermore, increasing the scale of the reaction resulted in the formation of by-products, complicating the isolation, separation and purification of the desired compound. We were however able to isolate pure samples of compound **27** and acquire the relevant analytical data which will be discussed in the following section.

The ^1H and ^{13}C NMR spectra of compound **27** can be found on the following page in **Figure 6.8**, indicating structural confirmation for the synthesis of the heterocyclic scaffold. Examination of the most upfield region of the ^1H spectrum in **Figure 6.8a** shows evidence of the propionamide moiety, with the terminal methyl (red) protons appearing as a triplet at δ 1.30 ppm and the neighbouring ethyl linker (green) as a quartet at δ 2.45 ppm. Moving downfield, the large peak found at δ 2.26 ppm was ascribed to the dimethylamine (light blue) element of the solubilising group. However, this signal was found to integrate for 8H instead of 6H, which was revealed to arise from overlap with another triplet upon closer inspection. This triplet was assigned to the protons of one of the ethylene bridge (dark blue) carbons, with the corresponding proximal triplet identified at δ 2.88 ppm. Lastly, the singlet appearing at δ 2.67 ppm and integrating for 3H was correlated to the methylamine (orange) of the solubilising group, expectedly more downfield than the aryl methoxy (yellow) moiety, appearing at δ 3.85 ppm.



Scheme 6.4: Coupling of fragments **14** and **26** to afford osimertinib-derived scaffold **27**. Reagents and conditions: i) TsOH·H₂O (1.2 equiv.), 2-pentanol, 105 °C, 8 h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745



Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

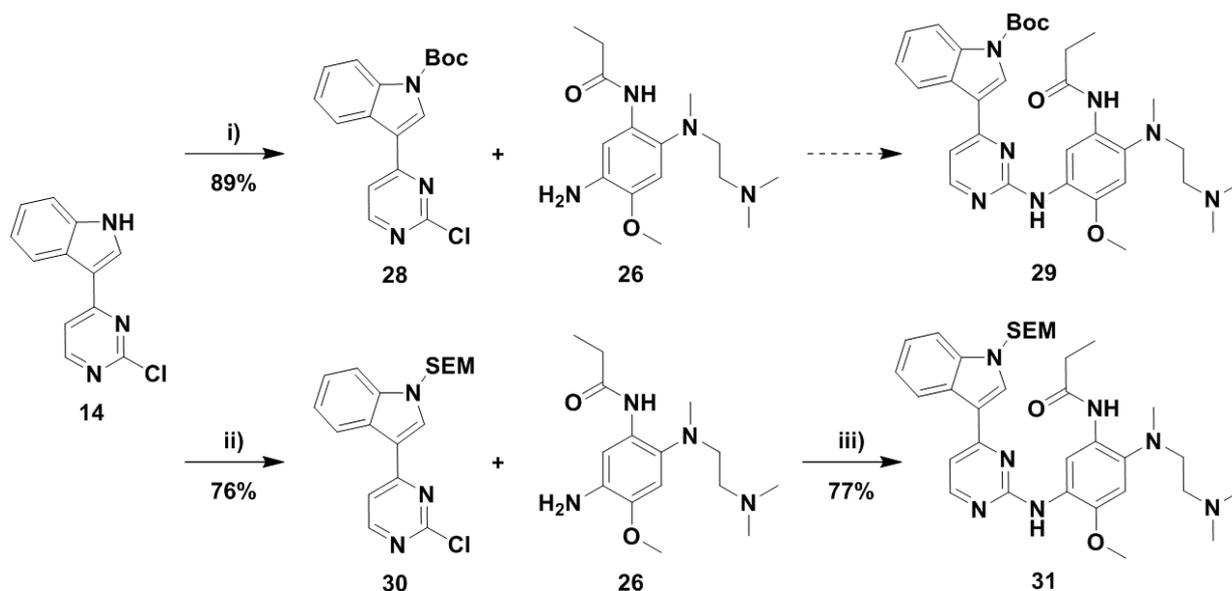
Assessment of the aromatic region of the ^1H NMR spectrum presented the correct integration amount for compound **27**, with all protons accounted for and peak splitting patterns to be in good order. This left us with three singlets at the farthest region downfield of the spectrum, corresponding to the remaining three amine protons in our structure. Working from NMR spectra of precursory structures and inherent differences in electron withdrawing substituents, we were able to assign the peak at δ 9.64 ppm to the aniline (purple) bridging the pyrimidine and aryl fragment, the propionamide (brown) at δ 9.97 ppm and lastly, the indole (grey) proton to the singlet at δ 10.38 ppm.

Inspection of the ^{13}C spectrum in **Figure 6.8b** further reinforces the structural findings of the ^1H spectrum, with several distinct peaks indicating the successful coupling of the two fragments. In the aliphatic region, the two peaks at δ 10.33 and 30.79 ppm correlate to the propionamide terminal methyl (green) and ethyl (red) carbons respectively. The remaining aliphatic region signals were assigned based upon visual comparison of the peak intensities and their position in the precursor fragments NMR spectra. Based on this, we assigned the most intense peak at δ 45.34 ppm to the dimethylamine (light blue) carbons and the methylamine (orange) carbon to the adjacent signal at δ 44.08 ppm. The three signals between δ 56.00 and 57.50 ppm proved more complex in their assignment, although we know they constituted the ethylene bridge (dark blue) carbons of the solubilising group and the aryl methoxy (yellow). Peak intensity, shape and previous spectra of isolated building blocks led us to assign the peaks as shown in **Figure 6.8b**, with the middle, more intense peak at δ 56.14 ppm belonging to the aryl methoxy (yellow) and the neighbouring peaks to the ethylene bridge (dark blue). At this point and with the information available, these assignments could not be confirmed and were considered speculative. However, 2D NMR spectroscopic analysis of a final compound with a similar osimertinib-derived heterocyclic driving group in the following chapter (**Section 7.4.3.9**) clarified these peak assignments.

To conclude, the carbons present in the aromatic region of the ^{13}C spectra were all correctly accounted for. The signal furthest downfield was indicative of a carbonyl carbon, which we ascribed to the propionamide (brown). Use of HRMS analysis presented a $[\text{M}+\text{H}]^+$ ion of 488.2783, with a calculated mass of 488.2774, providing final confirmation for the successful synthesis of compound **27**.

While pleased with the effective isolation and characterisation of compound **27**, we found the lower yields and inconvenient purification associated with scale-up to be unacceptable. We therefore sought an improved and more feasible reaction sequence for the synthesis of this compound. In the coupling between **14** and **26**, we suspected that the indole amine could be participating as a nucleophile owing to the high temperature of the reaction. We therefore proposed the introduction of protecting groups at this position to circumvent the unwanted formation of by-products, aspiring to improve the purification process and yield. As illustrated in **Scheme 6.4**, installation of a Boc-protecting group was carried out smoothly to afford compound **28** in excellent yield. Nonetheless, subsequent coupling between **28** and **26** led to the formation of numerous by-products, amongst which the deprotected compound **27** and the Boc-protected analogue **29** were identified by their isolation in trace amounts. We believed the requisite use of high temperatures and *p*-toluenesulfonic acid to be responsible for deprotection of the Boc-group, leading to several potential reactants and possible products. This prompted investigation into a more stable and chemically inert protecting group.

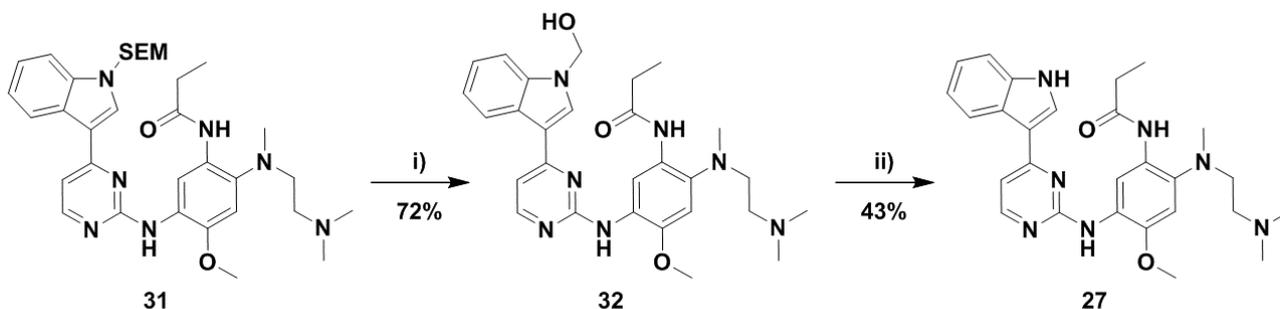
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745



Scheme 6.4: Introduction of protecting groups to improve synthesis of scaffold 27. Reagents and conditions: i) Boc₂O (1.2 equiv.), Et₃N (1.3 equiv.), DMAP (cat.), DCM, rt, 12 h; ii) NaH (1.16 equiv.), SEM-Cl (1.23 equiv.), -10 °C – rt, 4 h; iii) TsOH·H₂O (1.2 equiv.), 2-pentanol, 105 °C, 8 h.

This led us to the SEM-protecting group, with which we found marked success in the synthesis and protection of the pyrrolopyrimidine compounds of Chapter 3. Deprotonation of the indole amine with sodium hydride, followed by dropwise treatment with 2-(trimethylsilyl)ethoxymethyl chloride in DMF at low temperature, furnished the SEM-protected compound 30 in good yield. Fortunately, we found the coupling between 30 and 26 to proceed far more smoothly than in previous instances, with the SEM-protecting group surviving the harsh reaction conditions. Moreover, incorporation of the SEM-group resulted in the synthesis of “greasier” molecules with a lower polarity, facilitating easier handling and purification and resulting in an isolated yield of 77% of compound 31.

Deprotection of the SEM-group, shown in **Scheme 6.5**, proved unusually problematic. Unexpectedly, the previously effective established reaction conditions of trifluoroacetic acid in DCM, followed by 1 M KOH in THF, afforded the hydroxymethyl intermediate 32 almost exclusively amidst a few by-products. We therefore made use of trifluoroacetic acid in DCM independently to facilitate isolation of compound 32 in 72% yield. Following this, 32 was stirred in a 1:1 mixture of 1 M KOH and THF in a prolonged reaction which proceeded sluggishly and resulted in a maximum yield of 43% following challenging purification procedures.

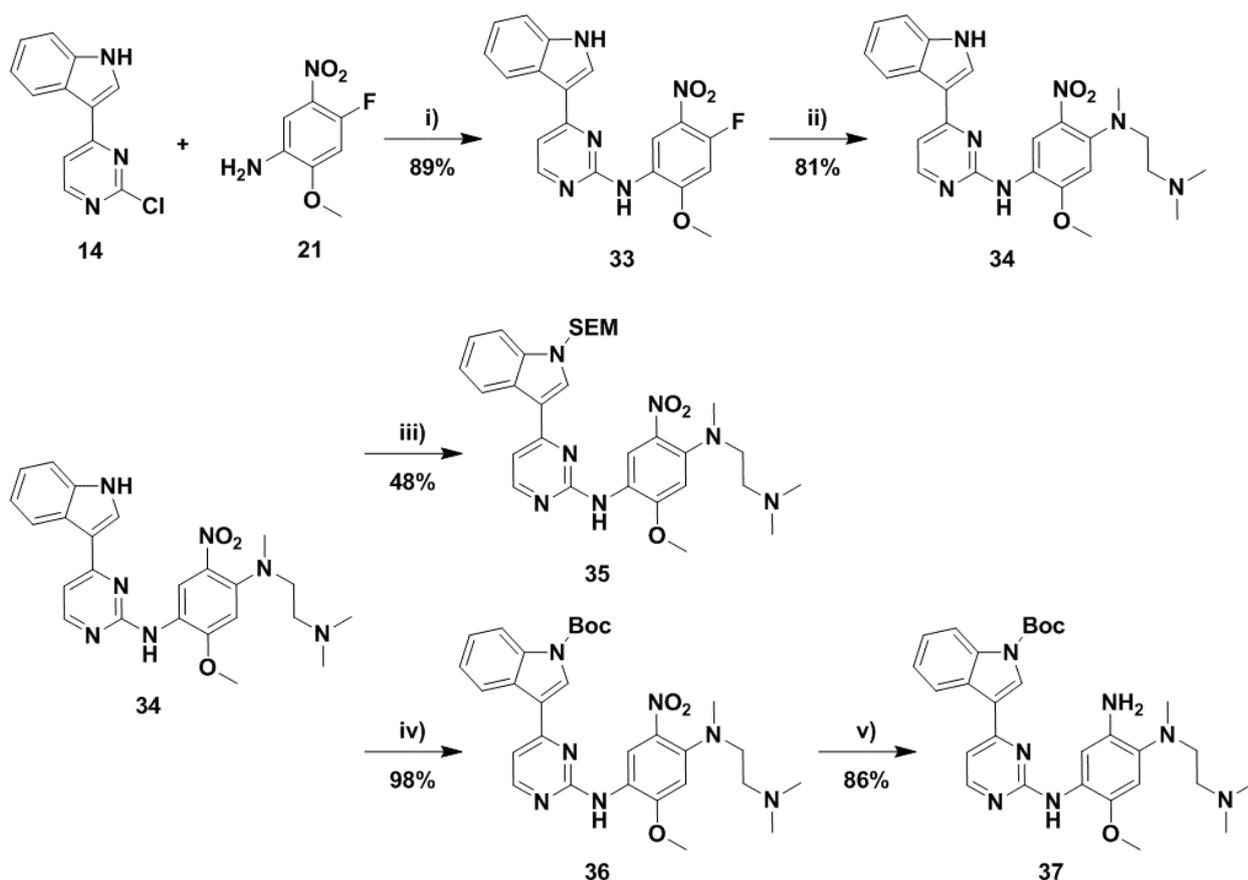


Scheme 6.5: Deprotection of SEM-protected compound 31 to afford intermediate 32 and scaffold 27. Reagents and conditions: i) DCM/TFA (3:1), rt, 6 h; ii) 1 M KOH/THF (1:1), rt, 4h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Assessment of the implemented synthetic strategy left much to be desired. The length of the pathway, requirement of protecting groups and incompatibility with reaction scale-up led to a final attempt towards route optimization, displayed below in **Scheme 6.6**. While coupling of fragments **14** and **21** and the stepwise building of the molecule in a linear fashion was originally contemplated, pursuit of this synthetic avenue posed several intrinsic chemo- and regioselective issues. The main concern was the inherent reactivity of the aryl fluorine centre of compound **21**, as exhibited previously in its Boc-protection to afford compound **22**. Exposing this fragment to high temperatures could result in multiple S_NAr reactions taking place, by nucleophilic attack of the indole amine or another equivalent of **21** to afford the dimerized product. While we hypothesised that this could potentially be overcome by prior introduction of the solubilising group, we nevertheless carried out the coupling of **14** and **21** using the optimized conditions.

To our surprise and delight, the reaction proceeded exceedingly well. This good fortune was augmented further by a serendipitous discovery regarding the purification of compound **33**. Contrary to the routine immediate purification of the crude reaction mixture, the reaction flask was left to cool overnight. This prompted the formation of a yellow precipitate, which was subsequently filtered and washed with acetonitrile. NMR spectroscopic analysis of this crude material revealed the presence of both the desired product **33** and excess *p*-toluenesulfonic acid.



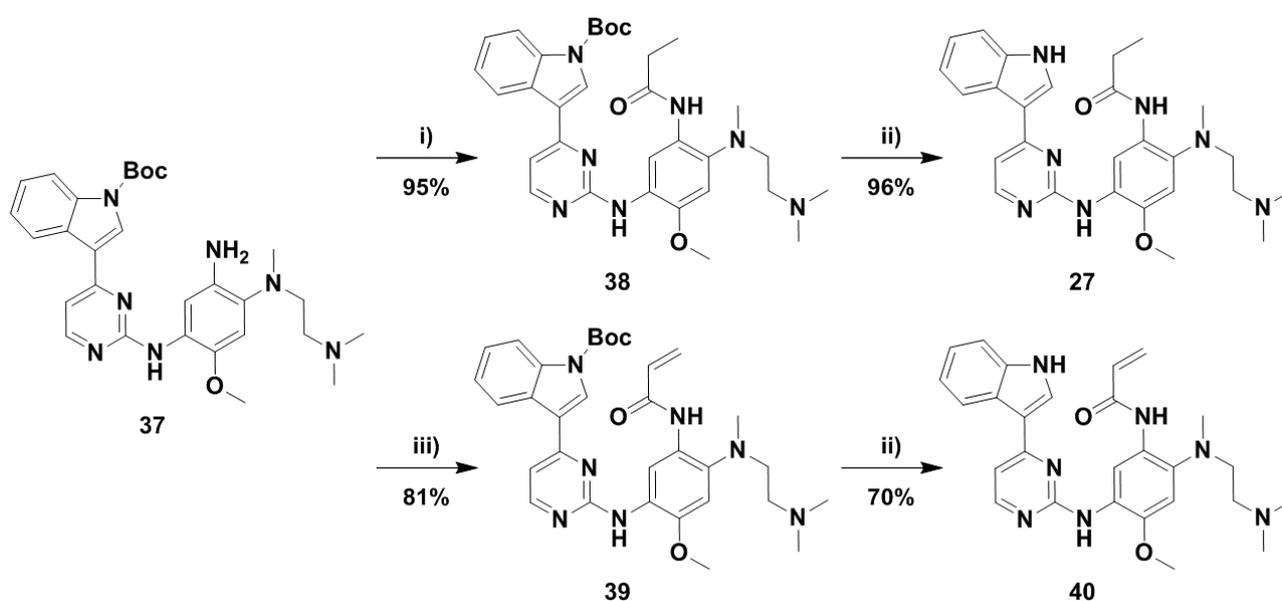
Scheme 6.6: Optimal synthetic route towards scaffold **27** by linear synthesis. Reagents and conditions: i) TsOH·H₂O (1.2 equiv.), 2-pentanol, 105 °C, 4 h; ii) *N,N,N'*-trimethylethane-1,2-diamine (1.25 equiv.), DIPEA (1.05 equiv.), DMA, 60 °C, 4 h; iii) NaH (1.2 equiv.), SEM-Cl (1.2 equiv.), -10 °C – rt, 4 h; iv) Boc₂O (1.1 equiv.), DIPEA (1.1 equiv.), DMAP (cat.), DMF, rt, 6 h; v) H₂, Pd/C (10 mol%), MeOH, rt, 12 h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Removal of the excess *p*-toluenesulfonic acid was achieved by dissolving the crude product in a 10% methanol in DCM solution, followed by successive washing of the organic phase with saturated sodium bicarbonate. The obtained yellow solid was then triturated with acetonitrile to afford compound **33** in 89% yield, with no further purification required. Exceptionally, this reaction could be scaled up to 6 grams with no repercussions to the obtained yield. Although the S_NAr mechanism remains a source of much investigation and debate, it is evident from the outcome of this reaction that protonation of the pyrimidine nitrogen, and thus facilitation of substitution at the centre *ortho* to this, is preferred.²⁶ In contrast, reaction at the fluorine centre is clearly favoured under basic conditions, as exemplified by the facile substitution and incorporation of the solubilising group in formation of **34**. Accordingly, mild heating of compound **33** and *N*¹,*N*¹,*N*²-trimethylethane-1,2-diamine in DMA, with DIPEA serving as base, afforded compound **34** as a bright red solid in good yield.

To avoid chemoselective complications between the indole amine and the aniline fragment in the forthcoming acylation reaction, we sought to introduce a protecting group at the indole nitrogen. Our first and singular attempt at the SEM-protection thereof resulted in a below average yield of 48% for compound **35**. Discouraged by this, we turned to the incorporation of a Boc-protecting group, using DMF as solvent for solubility purposes. The Boc-protection proceeded superbly, providing the protected analogue **36** in a near quantitative yield which was retained in subsequent scaled-up reactions. This allowed for subsequent aromatic nitro reduction using Pd/C under a hydrogen atmosphere to produce the aniline fragment **37** in excellent yield.

With **37** in hand, we advanced to the acylation using both propionyl and acryloyl chloride, as shown in **Scheme 6.7** below, which led to the formation of the penultimate protected compounds **38** and **39** respectively. For the installation of the Michael acceptor, yields were dramatically improved through implementation of low temperatures generated using an acetonitrile and dry ice bath.



Scheme 6.7: Incorporation of propionamide and acrylamide followed by deprotection to afford compound **27** and **40**.

Reagents and conditions: i) propionyl chloride (1.1 equiv.), DIPEA (1.2 equiv.), DCM, $-10\text{ }^{\circ}\text{C}$ – rt, 2 h; ii) DCM/TFA (3:1), rt, 4 h; iii) acryloyl chloride (1.1 equiv.), DIPEA (1.2 equiv.), DCM, $-40\text{ }^{\circ}\text{C}$ – $-20\text{ }^{\circ}\text{C}$, 2 h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Lastly, deprotection of the Boc-group was affected by stirring in a solution of trifluoroacetic acid in DCM. This afforded the desired scaffold **27** in near quantitative yield and the previously unattained Michael acceptor counterpart **40** in a decreased but acceptable yield of 70%. Development and use of this stepwise synthetic route resulted in an improvement of the overall yield, calculated to be 55.4% from the coupling of compound **14** and **21** to produce compound **27**. Moreover, the strategy was tolerant towards execution of multi-gram scale reactions, furnishing large amounts of scaffold **27** and **40** to carry forward in electrophile derivatisation. The synthesis of these electrophiles and the resultant final compounds will be disclosed in the following section.

6.4.2 Electrophile Synthesis and Derivatisation

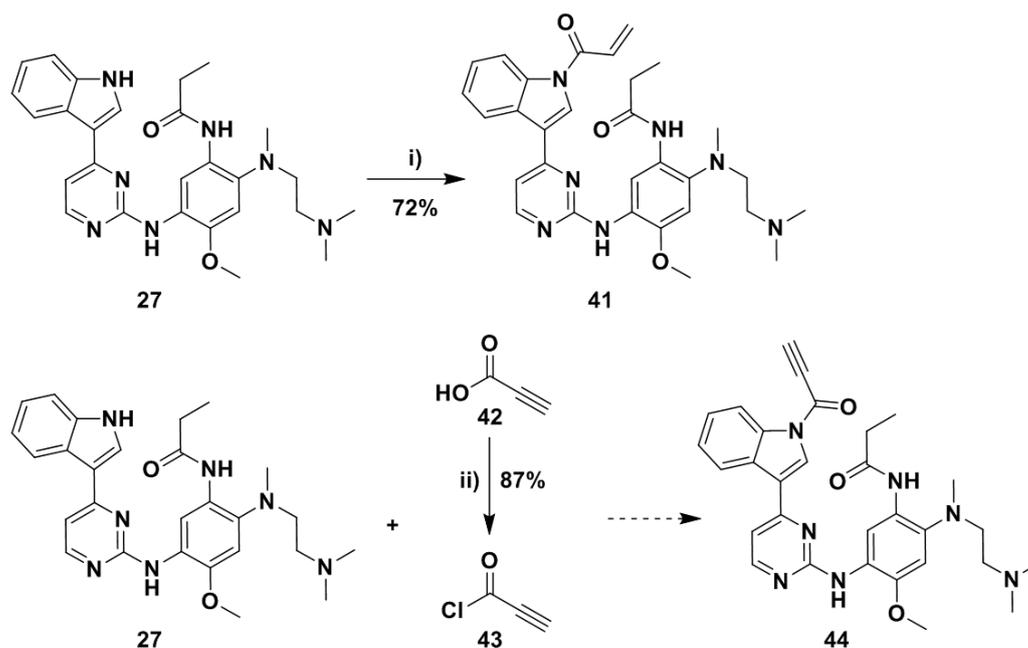
As the primary goal of this study was to gauge the feasibility of labelling the catalytic Lys745 residue, we chose to proceed exclusively with derivatization of the propionamide scaffold **27** in the generation of the first set compounds. This would eliminate any competition and avoid confusion as to whether covalent modification had occurred at the catalytic Lys745 or the traditional Cys797 residue. This would be relevant in covalent mass spectrometry experiments and the biochemical and cellular assays of enzyme variants other than the C797S mutant. Following evaluation of this preliminary library, we could decide whether to pursue the synthesis of compounds containing the Michael acceptor moiety derived from compound **40**.

6.4.2.1 Michael Acceptors

To our dismay, overall derivatization of the indole nitrogen proved extremely challenging, due largely to the unforeseen poor nucleophilicity of this functionality. Efforts at incorporating Michael acceptor electrophiles into scaffold **27** is illustrated in **Scheme 6.8** and the discussion thereof follows. Our first attempt, the seemingly straightforward acylation with acryloyl chloride using conventional conditions, resulted in no product formation and return of the starting material. A similar outcome was found in further experiments using an array of reaction parameters. This included the use of Et₃N, DIPEA or potassium carbonate as base at a range of reaction temperatures in various solvents such as DCM, THF and DMF. This was a source of much puzzlement and frustration, leading us to affect formal deprotonation of the indole nitrogen with sodium hydride. After being left to stir for 30 minutes at room temperature, the solution was cooled to -40 °C and treated dropwise with acryloyl chloride. Understandably, requirement of these harsh reaction conditions led to the formation of several side products. However, amidst these by-products, compound **41** was successfully isolated and afforded in a 72% yield, with DMF as optimal solvent.

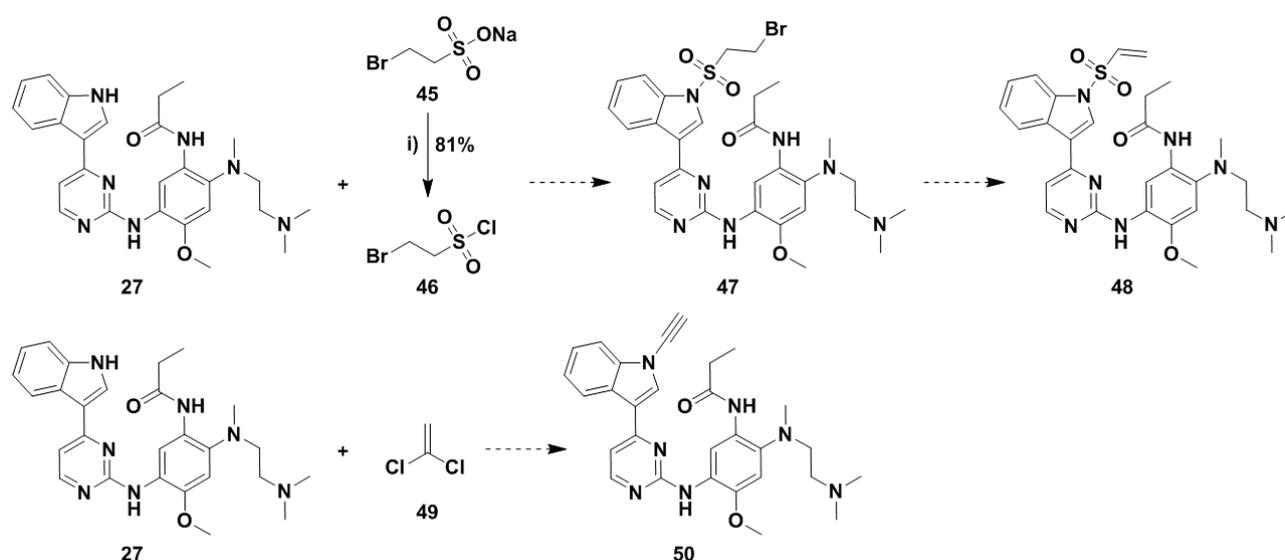
For inclusion of the propiolamide Michael acceptor, shown in **Scheme 6.8**, we had to contend with the aforementioned nucleophilicity issues. Initial failures towards amide bond formation using propiolic acid (**42**), led us to purify the corrosive liquid via kugelrohr distillation. However, similar results were obtained in subsequent attempted acid couplings, using an assortment of reagents such as DCC, EDC, DIC and CDI in combination with HOBt hydrate and DMAP as additives. Hoping to emulate the workable conditions achieved with acryloyl chloride, conversion of **42** into the corresponding acid chloride **43** was successfully carried out using phosphorous pentachloride in DCM. Unfortunately, use of sodium hydride and subsequent treatment with **43** led to a myriad of side reactions, making it near impossible to distinguish and isolate **44**.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745



Scheme 6.8: Synthesis of the Michael acceptor-containing compound **41** and the attempted synthesis of final compounds **44**. Reagents and conditions: i) NaH (1.3 equiv.), $-10\text{ }^{\circ}\text{C}$ – rt, 30 min then acryloyl chloride (1.2 equiv.), DMF, $-40\text{ }^{\circ}\text{C}$ – $20\text{ }^{\circ}\text{C}$, 2h; ii) PCl_5 (1.05 equiv.), DCM, rt, 4h.²⁷

We shifted our focus to the devised synthetic strategy outlined in **Scheme 6.9**. To obtain ethenesulfonamide **48**, acylation of **27** with the synthesised **46** would be undertaken, followed by elimination of **47**. Accordingly, neat reaction of the commercially available sodium salt of bromoethansulfonate (**45**) with phosphorous pentachloride under thermal conditions provided **46** in good yield.²⁸ Regrettably, compound **47**, and therefore compound **48**, could not be successfully isolated. Once again, acylation of the weakly nucleophilic indole nitrogen required strong basic conditions to facilitate any reaction, resulting in the formation of numerous side-products and a substantial remainder of starting material.



Scheme 6.9: Synthesis of reactive intermediate **46** and the attempted synthesis of final compounds **48** and **50**. Reagents and conditions: i) PCl_5 (1 equiv.), neat, $130\text{ }^{\circ}\text{C}$, 30 min.^{28, 29}

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

We believe these difficulties stem from an inherent catch-22 situation, where requisite use of a strong base induces spontaneous elimination of **46** to form the ethenesulfonyl moiety *in situ*, mainly attributed to the highly reactive bromo-leaving group. This could occur prior to or post acylation and generates multiple electrophilic sites, evinced by the observed mass of by-products. While unsuccessful in the synthesis of compound **48**, it is worthy to note that this strategy was successfully implemented in the generation of a different compound, where use of strong base was not obligatory. The synthesis of this compound will be discussed in **Section 7.4.3.1** in the following chapter.

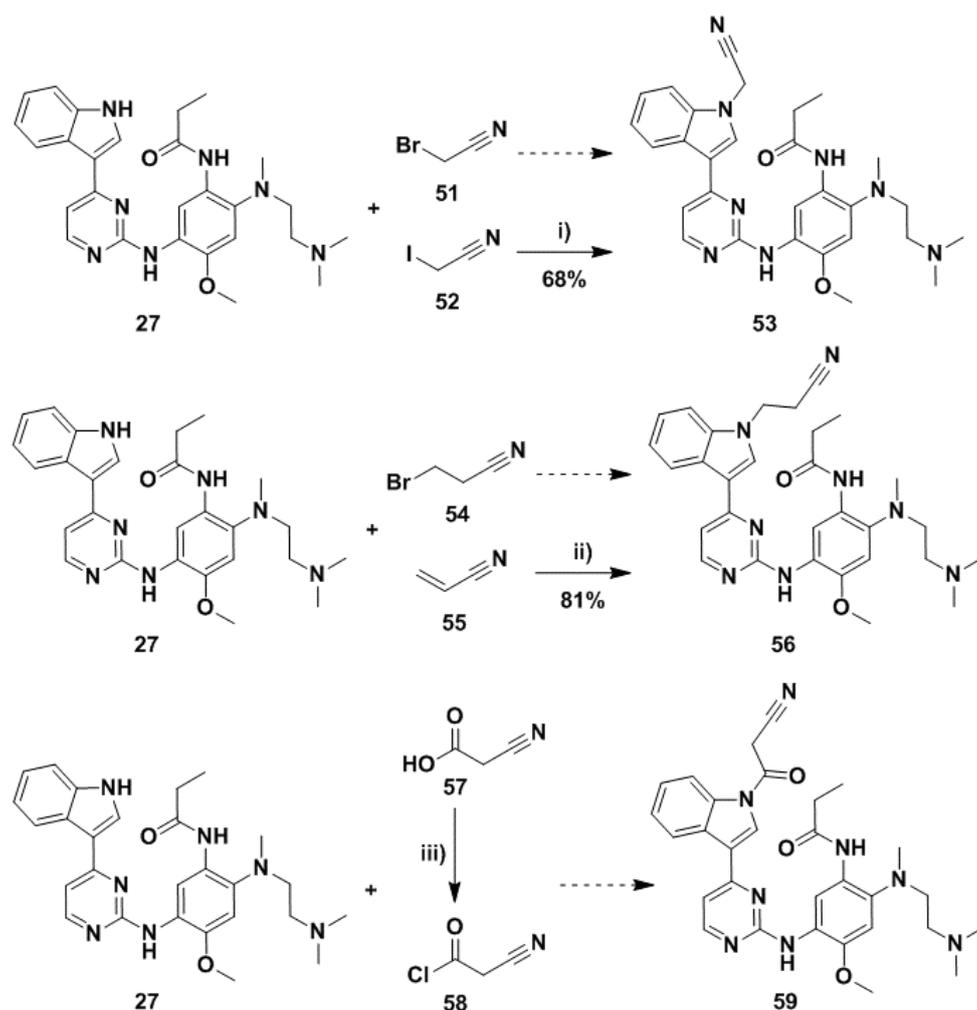
To overcome this dilemma, we envisioned use of the commercially available chloro-variant of **47**, which contained a weaker leaving group and would thus be less prone to undergo the untimely elimination. Unfortunately, due to the reactive and unstable nature of the compound, the danger posed to its transport and delivery inflated the cost to exorbitant amounts. Additionally, shipping times would be in the vicinity of 6 – 8 months, leading us to suspend efforts towards the synthesis of compound **48**. However, prospective routes towards the synthesis of **48** can be found in the future work section of this chapter. Due to similar constraints, the inability to access commercially available electrophiles required for the synthesis of final compounds in this and the following chapter became a recurring trend. This meant that these reactive fragments had to be synthesised in house, adding a further layer of complexity and synthetic labour to the research project.

In recent years, ynamides have found increased use as Michael acceptors in the development of covalent kinase inhibitors that target the traditional cysteine residue.¹⁸ We therefore attempted the synthesis of ynamide **50**, shown in **Scheme 6.9**, using the robust protocol developed by Tu and co-workers.²⁹ The group found that following reaction of amines with 1,1-dichloroethylene (**49**), the masked alkynyl halide could be transformed further under basic conditions into the corresponding ynamide using a broad substrate scope. However, the prescribed use of 5 equivalents of sodium hydride in heated DMSO was simply too harsh for our heterocyclic scaffold. This led to the now familiar formation of inseparable crude reaction mixtures and consequent abandonment of the synthesis of compound **50**.

6.4.2.2 Nitriles

Equipped with a greater understanding of the indole nitrogen reactivity, we were pleased to find that introduction of the nitrile functional groups proved simpler and more fruitful. As illustrated in **Scheme 6.10**, initial reaction between **27** and bromoacetonitrile (**51**) using sodium hydride as base was met with no product formation and a return of the starting material. Accordingly, exchange of electrophile to the highly reactive iodoacetonitrile (**52**) and application of the same conditions rewarded us with the efficient production of compound **53** in 68% yield.

Comparatively, the synthesis of **56** was met with failure when reacting 3-proprionitrile (**54**) with **27** using a variety of bases and Finkelstein type conditions at elevated temperatures. Interestingly, success was attained through use of acrylonitrile (**55**) and benzyltrimethylammonium hydroxide (Triton B), a quaternary salt that can act as an effective organic base catalyst in Michael addition reactions.³⁰ Consistent with the literature, the addition proceeded smoothly at ambient temperature over a slightly longer reaction time, furnishing **56** in excellent yield. While not discussed, both compound **53** and **56** underwent full characterization which may be accessed in the supplementary information.



Scheme 6.10: Synthesis of nitrile-containing final compounds **53** and **56** and the attempted synthesis of compound **59**.

Reagents and conditions: i) NaH (1.3 equiv.), $-10\text{ }^{\circ}\text{C}$ – rt, 30 min then iodoacetonitrile (1.2 equiv.), DMF, $-10\text{ }^{\circ}\text{C}$ – rt, 2 h; ii) Triton B (cat.), acrylonitrile (4 equiv.), DMF, $-10\text{ }^{\circ}\text{C}$ – rt, 24 h; iii) oxalyl chloride (1.2 equiv.), DMF (cat.), DCM, $40\text{ }^{\circ}\text{C}$, 4 h.³⁰

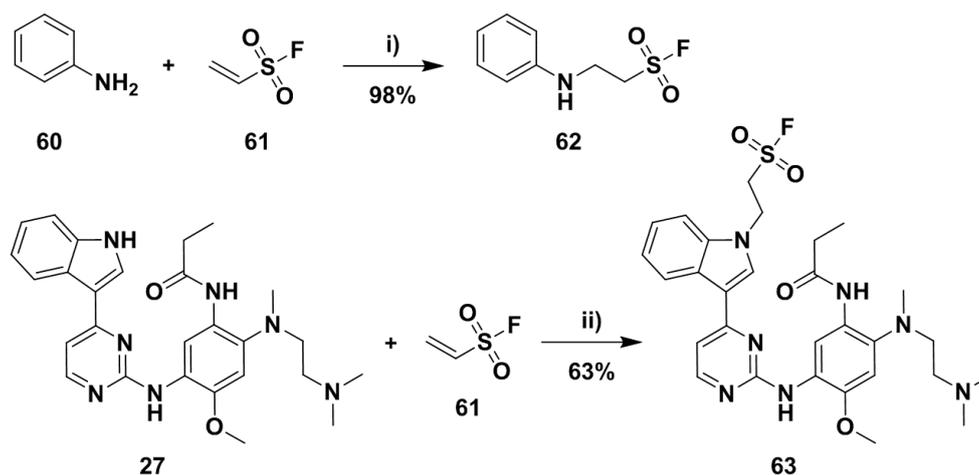
31

Finally, venture into the incorporation of a cyanoacetamide moiety into structure **27**, to potentially afford compound **59**, is shown above in **Scheme 6.10**. The attempted synthesis of **59** was carried out in a manner comparable to that undertaken for propiolamide **44** in **Scheme 5.8**. As such, the synthesis fell prey to the same pitfalls discussed for compound **44** and we were unable to successfully isolate compound **60**, even from reaction of **27** with the synthesised 2-cyanoacetyl chloride (**58**). It is noteworthy that execution of this strategy with a more nucleophilic aniline fragment was met with success and will be reviewed in **Section 7.4.3.8** in the following chapter.

6.4.2.3 Sulfonyl Fluorides

Until recently, ethenesulfonyl fluoride (ESF) and its use as a reactive building block remained relatively obscure. Shown in **Scheme 6.11** as compound **61**, ESF was first reported by Hedrick in 1953 and its reactivity systematically investigated by Krutak and co-workers in 1979.³² By virtue of its highly reactive C=C bond, which has been distinguished as one of the most electrophilic monosubstituted ethylenes, reaction with a variety of nucleophiles is facilitated whilst sparing the sulfonyl fluoride moiety which remains unaffected.³³

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745



Scheme 6.11: Synthesis of model compound **62** and its final compound counterpart **63**. Reagents and conditions: i) ESF (1.05 equiv.), DMF, 0 °C – rt, 3 h; ii) ESF (1.3 equiv.), K₂CO₃ (1.5 equiv.), DMF, 0 °C – rt, 6 h.³²

ESF has found numerous applications in cross-coupling and click chemistry, leading to it being named “the most perfect Michael acceptor ever found,” by Sharpless and his research group.^{6, 34} In light of our inexperience in handling sulfonyl fluoride-containing molecules, we undertook a preliminary model reaction between aniline (**60**) and ESF (**61**), illustrated in **Scheme 6.11** above. By doing so, we wished to gain a better understanding of the functional group’s reactivity, sensitivity and stability towards purification procedures such as aqueous work-up and column chromatography. It was found that dropwise treatment of **60** with diluted ESF (**61**) at low temperatures in DMF facilitated the Michael addition effortlessly. Aqueous work-up was well tolerated and extraction with ethyl acetate ensured removal of any excess DMF. This was followed by purification using column chromatography which afforded **62** in a near quantitative yield, highlighting the facile conversion and stability of the sulfonyl fluoride.

With our curiosity and apprehension towards the reactive group allayed, we applied the same reaction conditions to the synthesis of compound **63**. While this resulted in formation of **63**, the reaction proceeded sluggishly with limited yields. Inclusion of potassium carbonate in the reaction, serving as base and driving the Michael addition forward, gave rise to improved yields and reaction times. Consequently, we were able to isolate final compound **63** in an acceptable yield of 63%.

With both the model compound **62** and final compound **63** in hand, we were able to perform comparative spectroscopic analysis using ¹H and ¹⁹F NMR spectroscopy. These spectra are shown on the following page in **Figure 6.9** and were used to provide confirmation for the successful synthesis of **63**. All ¹⁹F NMR spectroscopy experiments were self-run, with chemical shifts within spectra referenced to trichlorofluoromethane which served as the internal standard. Inspection of the ¹⁹F NMR spectra of **62**, found in the top right of **Figure 6.9a**, displays a singular peak (red) at δ 56.70 ppm which coincides with the lone fluorine atom of the sulfonyl fluoride (red). Interestingly, this peak was revealed as a triplet coupling through the sulfonyl moiety to the neighbouring protons of the ethyl chain linker, with a J-coupling value of 4.6 Hz. Combing through the literature, we found similar compounds whose ¹⁹F NMR shifts coincides with this peak (~57 ppm) and corroborated with the observed multiplicity and J-coupling value (~ 4.8 Hz).³⁵ Interestingly, ¹⁹F NMR spectroscopy reporting for this class of compound is very sparse, with only several examples found in publications spanning the last two years.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

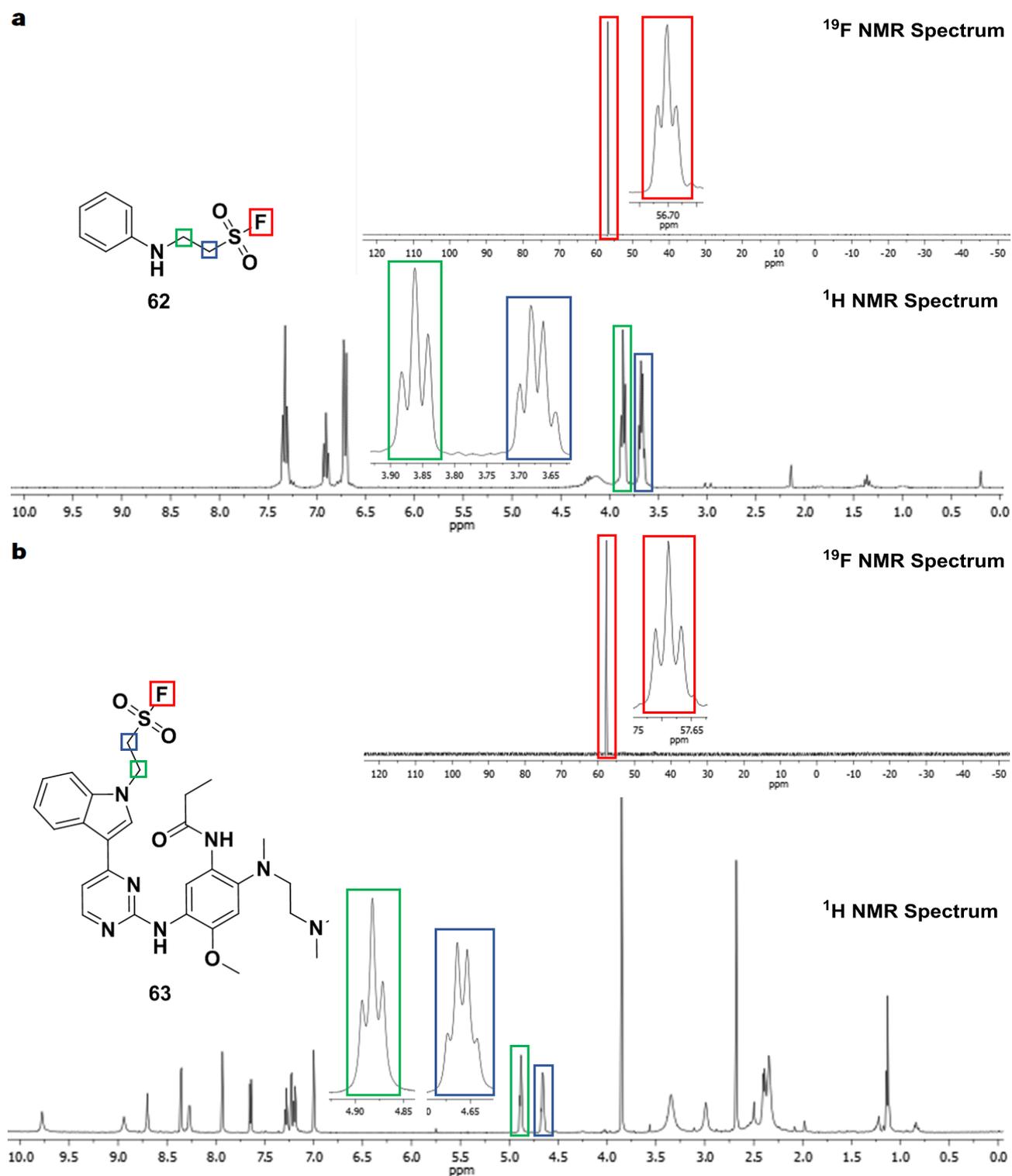


Figure 6.9: ^1H and ^{19}F NMR of a) compound **62** and b) final compound **63**.

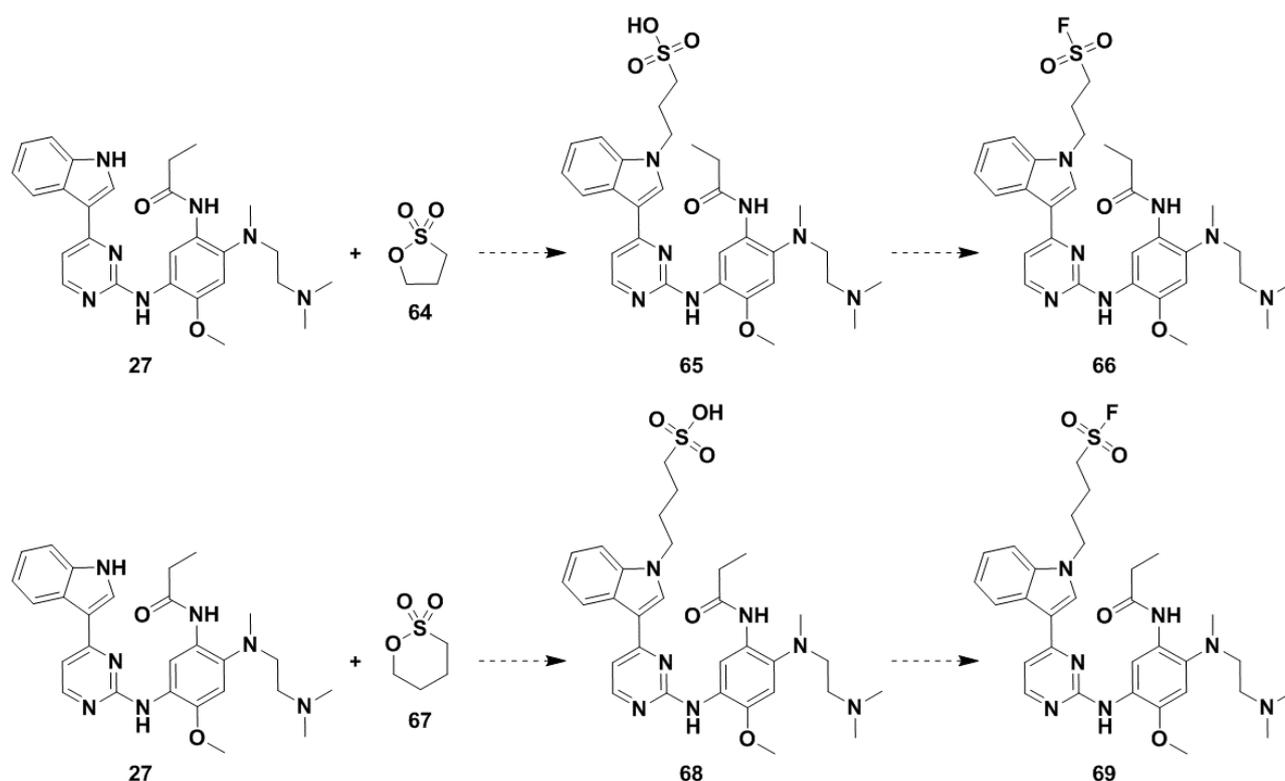
From the ^1H NMR spectra we gained confirmation of the synthesis of compound **62**, with all protons correctly accounted for. Importantly, the presence of the two peaks (blue and green) between δ 3.60 and 3.90 ppm were ascribed to the ethylene bridge of the sulfonyl fluoride. The most upfield peak at δ 3.67 ppm, which appeared as an apparent quartet, was assigned to the methylene (blue) closest to the sulfonyl group. We believe this multiplicity arose from coupling to the neighbouring methylene and long-range coupling through the sulfonyl group to the fluorine atom, resulting in the overlap of a triplet of doublets and the observed quartet.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

This left the remaining bridge methylene (green) signal, which occurred slightly downfield at δ 3.86 ppm as an apparent triplet.

With the model compound fully characterised, we undertook a comparative analysis between final compound **63**. Examination of the ^{19}F NMR spectrum of **63**, shown in the top right inset of **Figure 6.9b**, revealed a striking similarity to that of **62**. A single peak was observed at δ 57.69 ppm, less than 1 ppm from the model compound signal and in the vicinity of similar compounds found in the literature. Furthermore, both peaks appeared as triplets, with **63** exhibiting a slightly higher J-coupling value of 6.1 Hz. Successful synthesis of the sulfonyl fluoride containing final compound **63** was endorsed further by the obtained ^1H NMR spectra. As with the model compound, the emergence of two new signals between δ 4.60 and 5.00 ppm, correlating to the ethylene bridge of the sulfonyl fluoride, was witnessed. These peaks (blue and green), while slightly downfield from the model counterpart, retained the same multiplicity and analogous J-coupling values, substantiating successful sulfonyl fluoride functionalisation. All other protons, including those of amino, aliphatic and aromatic, were correctly accounted for. Except for the absence of the indole nitrogen proton, the ^1H NMR spectra of **63** compares well with the precursor scaffold **27** in **Figure 6.8** and will therefore not be discussed further. Lastly, HRMS analysis provided an $[\text{M}+\text{H}]^+$ ion of 598.2620, matching the calculated mass of 598.2612, providing final confirmation for the successful synthesis of compound **63**.

To diversify our library of sulfonyl fluoride compounds, we wished to synthesise the propyl (**66**) and butyl (**69**) bridge counterparts of **63**, illustrated in **Scheme 6.12** below. To accomplish this, we envisioned ring-opening reaction between scaffold **27** and both the 1,3-propane and 1,4-butane sultone **64** and **67** respectively. This would afford the sulfonate intermediates **65** and **68**, which could then be converted into the corresponding sulfonyl fluorides.



Scheme 6.12: Attempted synthesis of final compounds **66** and **69**.

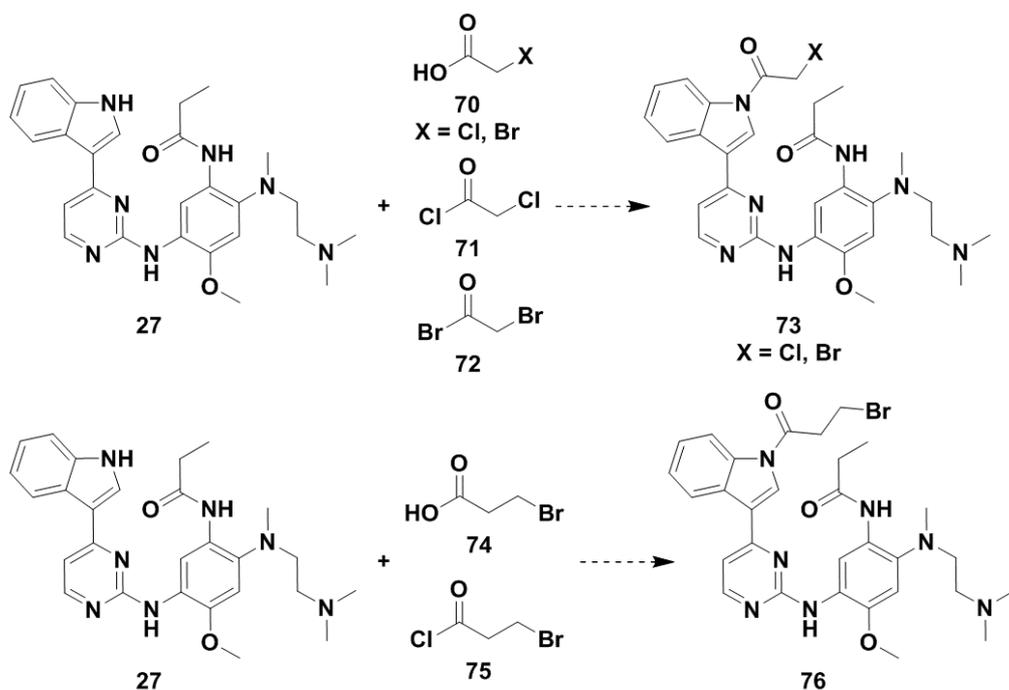
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

While we believe the ring-opening reaction was affected successfully using sodium hydride and heating in DMF, purification and characterization of the intermediate sulfonate compounds **65** and **68** proved impractical. These compounds were found to be extremely polar and water-soluble, leaving all products, by-products and excess starting materials within the aqueous phase following work-up, as visualised by TLC analysis. Subsequent removal of water using a lyophilizer afforded a crude mass which was subjected to NMR spectroscopic analysis. While it could be surmised that the reaction was successful, this could not be confirmed as the obtained spectra was largely indecipherable due to the presence of numerous by-products. Furthermore, purification of the crude mixture proved futile due to the compound's solubility and polarity.

Nevertheless, conversion of the crude sulfonates **65** and **68** to the corresponding sulfonyl fluorides **66** and **69** was attempted using several reagents and reaction conditions. This included use of the fluorinating reagent diethylaminosulfur trifluoride (DAST), conversion to the equivalent sulfonyl chloride using oxalyl chloride followed by substitution with potassium fluoride and numerous one-pot syntheses, all to no avail. We believe implementation of reverse-phase column chromatography to isolate pure samples of the sulfonate intermediates **65** and **68** will allow for the successful synthesis of final compounds **66** and **69**. Unfortunately, this purification technique was unavailable to us, but it remains our endeavour to synthesise these compounds as part of future work of this chapter.

6.4.2.4 α -, β -Haloketones

Efforts towards the synthesis of α - and β -haloketone-containing final compounds, depicted in **Scheme 6.13** below, will be discussed briefly. For the synthesis of both the halide-variants of **73** and **76**, acid coupling and acylation reactions were attempted using a variety of carboxylic acid and acyl halide derivatives.



Scheme 6.13: Attempted synthesis of final compounds **73** and **76**.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

All of the electrophiles employed in these reactions were distilled or recrystallised prior to use. Unfortunately, no product formation occurred for the acid coupling of **27** with either **70** or **74**, owing to the inherent weak nucleophilicity of the indole nitrogen. However, for the duration of acylation reactions carried out at low temperatures, it appeared as though single product formation was taking place, as per visualisation by TLC analysis. However, following the attempted isolation and purification of the crude reaction mixtures, the compounds were found to have degraded, evident by the abundant formation of side-products. This led to our inability to furnish compounds **73** and **76**.

In the following chapter, similar reactions were carried using a stronger nucleophilic aniline fragment. However, these compounds also succumbed to the same outcomes of side-product formation and degradation. Our theories towards the reactivity and instability of these compounds will be broached in the next chapter and a full explanation, with compelling evidence provided by by-product formation, will be given in **Section 7.4.3.6**.

As we wished to ascertain whether our strategy was viable in targeting interactions with the catalytic Lys745, it was decided to evaluate the inhibitory properties of the successfully synthesised final compounds until this point. This concluded the synthesis of our osimertinib derived lysine-targeting compounds and allowed for the biochemical and biological assessment of all final compounds and interesting intermediates. These compounds were tested against the wild-type and mutant variant forms of EGFR, the results of which are discussed in the following section.

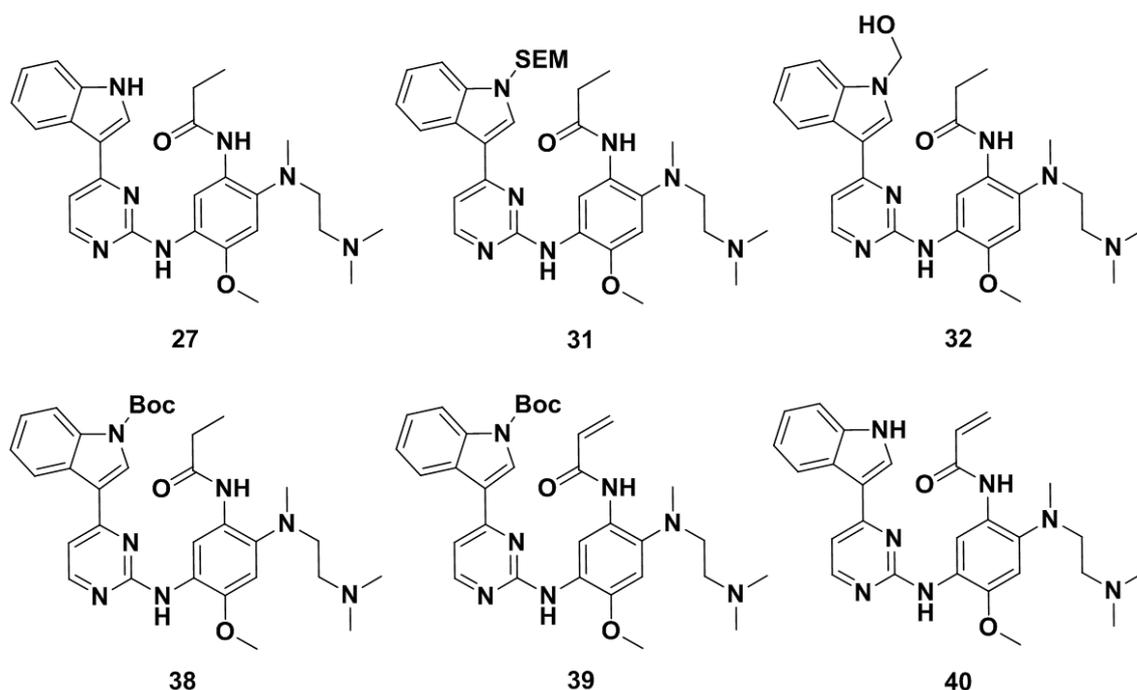
6.5 Biochemical and Cellular Evaluation

Unfortunately, complications within the biology lab of our collaborators at the Technische Universität in Dortmund prevented biochemical inhibitory measurements of the synthesized final compounds against the triple L858R/T790M/C797S mutant variant of the enzyme. However, we were able to acquire the EC₅₀ values against the clinically relevant osimertinib resistant cell line EGFR-PC9/T790M/C797S in the cellular evaluation, with the PC9 mutation being the predominant activating mutation found in osimertinib resistant patients. Furthermore, a wild-type-bearing cell line A431 and the EGFR-L858R/T790M drug resistant cell line H1975 was employed. Additionally, promising compounds were further evaluated against BaF3 cell lines of both the single activating EGFR-L858R and the EGFR-T790M/C797S mutant variants of the enzyme (**Section 4.6.3**). Compound assessment was carried out either by Dr. Jonas Lategahn, Dr. Marina Keul or Mr. Tobias Grabe, with all measurements being executed in at least triplicate. The respective IC₅₀ and EC₅₀ values of osimertinib and the underivatized analogue **27** accompany the obtained results in each report.

6.5.1 Evaluation of Intermediate Compounds

An overview of the results from the biochemical and cellular evaluation of all intermediate compounds synthesised en route to the final compounds of the chapter can be seen in **Table 6.1** on the following page. As these precursors are derivatised at the *N*-indole position with a SEM- or Boc-protecting group, we wished to investigate any potential interactions that the silyl ether or carbamate may have with the Lys795 residue and the optimal chain length for derivatisation. These compounds could then be compared to the underivatized analogue **27**, whose entry is found just below osimertinib in **Table 6.1**.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Table 6.1: Biochemical and cellular evaluation results for the synthesised intermediate compounds.

Compound	EGFR HTRF IC ₅₀ [nM]			EGFR CTG EC ₅₀ [nM]				
	wild-type	L858R	L858R/ T790M	A431	H1975	PC9/ T790M/ C797S	BaF3/ L858R	BaF3 T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	756 ± 340	16.0 ± 5.00	3109 ± 648	< 14	2041 ± 590
27	75.9 ± 12.8	15.7 ± 1.33	19.0 ± 13.1	7981 ± 3270	7653 ± 850	6880 ± 2550	-	-
31	747 ± 159	86.1 ± 4.55	127 ± 9.39	8933 ± 1230	9351 ± 2870	3828 ± 210	< 14	25010 ± 8470
32	140 ± 42.4	26.7 ± 2.69	34.2 ± 1.16	9548 ± 710	7904 ± 1900	8143 ± 660	-	-
38	1219 ± 115	221 ± 34.1	222 ± 22.8	30000 ± 0	16500 ± 1909	30000 ± 0	-	-
39	0.32 ± 0.02	0.15 ± 0.01	0.08 ± 0.00	930 ± 518	90.0 ± 12.3	2568 ± 540	< 14	3000 ± 130

Analysis of the biochemical data for model compound **27** reveals a reasonable activity profile, with understandable loss of efficacy when compared to osimertinib, owing to the absence of the Michael acceptor and covalent bond formation with Cys797. Surprisingly, the IC₅₀ values against the single activating L858R and double L858R/T790M mutant variants remains within the double digit nanomolar range. However, analysis of the corresponding cellular data reveals the significance of exclusion of the Michael acceptor moiety and loss of covalent bond formation, with a dramatic loss in activity across the board. We see low nanomolar inhibitory concentrations in the biochemical assessment translating to high single digit micromolar values in the cellular assays, highlighting the potentially misleading presentation of biochemical data and the importance of attaining good cellular activity as a benchmark for further investigation.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Scrutiny of the SEM-protected and partially deprotected compounds **31** and **32** respectively, reveals interesting trends regarding chain length and potential metabolite formation. Biochemically, compound **32** globally outperforms the SEM-protected counterpart **31**, with a 3- to 5-fold improvement in IC_{50} values. This alludes to the shorter hydroxymethyl chain of **32** being superior to the longer trimethylsilylethoxymethyl chain of the SEM-protecting group of **31**. Furthermore, the more polar terminal hydroxyl of **32** may induce hydrogen bond interactions with the polar residues of the $\beta 3$ sheet or helix αC strand at the roof of the enzyme, where the trimethylsilyl ether of **31** cannot. This matches with the hydrogen bond donor capabilities of the indole nitrogen of model compound **27**, which exhibits slightly improved IC_{50} values over compound **32**.

As witnessed in compound **27**, we see a large diminishment of activity in the EC_{50} values of the both **31** and **32**, ascribed to the absence of covalent modification of the cysteine residue. Interestingly, compound **31** displays an improved cellular activity against the PC9/T790M/C797S triple mutant, surpassing both the model compound **27** and **32**. As such, compound **32** was evaluated against BaF3 cell lines of both the single activating L858R and the T790M/C797S EGFR mutants. As mentioned in Chapter 3, the use of Ba/F3 cells in chemical biology has shown promise in predicting resistance by point mutations which interfere with inhibitor binding, allowing for an estimation of the candidate compounds susceptibility to clinical resistance.³⁶ While compound **31** retained efficacy comparable to osimertinib against the BaF3/L858R cell line, a complete loss in activity was observed against BaF3/T790M/C797S, with an EC_{50} value of 25 μM .

Next, we inspected the performance of the Boc-protected compound **38** and **39**. The propionamide derivative **38** was shown to be universally less efficacious than the SEM-protected analogues biochemically and essentially inactive in the cellular screening. This strengthened our hypothesis of shorter chain length and a polarised terminal functionality, as the carbamate chain is both long and contains a *tert*-butyl end group. Intriguingly, the unsaturated counterpart **39** displayed incredible efficacy biochemically, outperforming both the model compound **27** and osimertinib with IC_{50} values in the picomolar range. Compound **39** exhibited an IC_{50} value of 80 pM against the L858R/T790M double mutant, nearly 7-fold more potent than the structurally related osimertinib. With the presence of the Michael acceptor restored, these values were retained in the cellular assays. Double digit nanomolar activity against the L858R/T790M and a slightly improved EC_{50} value of 2.57 μM for the triple mutant was observed. Excited by this, compound **39** was run against the BaF3 cell lines which revealed an activity profile comparable to osimertinib.

Seeking an explanation for this remarkable activity, we scoured the literature for conceivable answers. Reading through the original manuscript describing the development of osimertinib, we came across the evaluation of compound **40** (Table 6.1), which we also synthesised but did not screen.³⁷ According to the authors and following *in vivo* metabolic studies in murine plasma samples, this compound is a metabolite of osimertinib, resultant from loss of the indole *N*-methyl group. While profiling of compound **40** showed impressive double and activating mutant potency, the candidate was not developed further due to the modest *in vitro* wild-type margin and *in vivo* efficacy. As compound **39** displays a similar inhibitory profile, we propose that the Boc-protecting group is being cleaved during incubation, potentially resulting in formation of metabolite **40**. Additionally, this theorem correlates with the observed superior activity of model compound **27**, which is also unsubstituted at the indole nitrogen.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

From a drug-delivery perspective, we believe there is potential in the further development of similar compounds containing protecting groups which are readily metabolizable. This concluded the evaluation of the precursor compounds, with discussion of the final compounds inhibitory performance to follow.

6.5.2 Evaluation of Final Compounds

Final compounds bearing electrophiles targeting the catalytic Lys745 residue were screened both biochemically and in a cellular setting, the results of which may be found in **Table 6.2**. Foremost inspection of the IC_{50} values associated with the indole acrylamide-substituted **41** showed a general loss in activity when compared with model compound **27**. These unimpressive values were reflected in the cellular data, with high micromolar EC_{50} values indicating intolerance and a lack of favourable interactions generated by acryloyl derivatisation at this position.

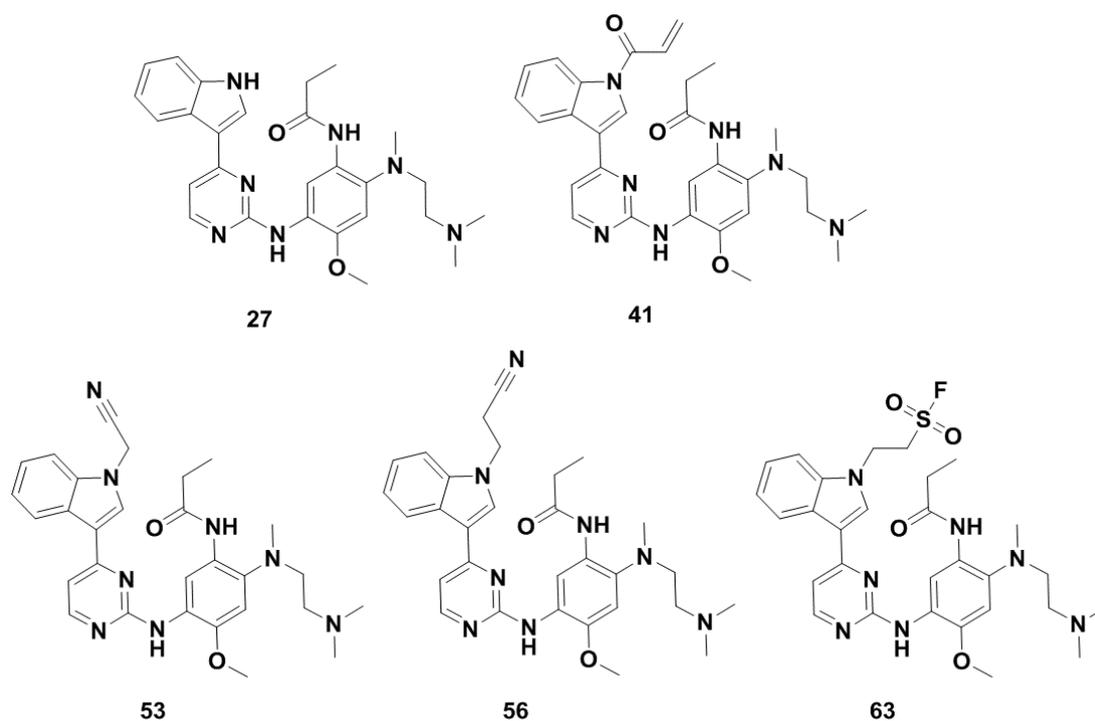
Proceeding to the nitrile derivatives **53** and **56**, we witnessed improved inhibition values and several noteworthy trends. Comparative analysis of the biochemical data between these two compounds revealed a superior activity profile for **56**, with enhanced activity over **27** against both the activating and double mutant. Furthermore, this compound displayed a favourable wild-type selectivity margin, exhibiting 8- and 10-fold selectivity for the L858R and L858R/T790M mutant variants respectively. We believe this enhanced activity could stem from the nitrile chain, with compound **56** being lengthened by an extra carbon and potentially allowing for reversible interactions of the terminal nitrile with the catalytic Lys795 or polar residues at the roof of the kinase.

Cellular assessment of compound **53** presented EC_{50} values that were nearly identical to **27**, barring a slight improvement against the PC9/T790M/C797S mutant. Interestingly, compound **56** retained its positive selectivity profile and displayed the greatest inhibition of the triple mutant. This warranted an examination of the BaF3 cell lines, where compound **56** excelled over its shorter chain counterpart with an over 2-fold improved activity of $5.70 \mu M$. Although this value is inferior to that of osimertinib, we still maintain that use of the long chain nitrile functionality of compound **56** bears potential in the development of future drug candidates.

This brought us to the final compound of the series, the sulfonyl fluoride containing **63**. To our disappointment, compound **63** was found to be completely inactive in the cellular assays, reaching the EC_{50} value limit of $30.0 \mu M$. Consequently, our collaborators who undertook the screening chose not to submit this compound for biochemical evaluation. It is our belief that this inhibitor either did not survive the journey from South Africa to Germany, which could have potentially occurred through compound degradation or hydrolysis or the inhibitor was being hydrolysed to the corresponding sulfonic acid within the cells during their assessment. As such, we still intend to resynthesise and dispatch this compound to our associates, along with another derivative which will be discussed in the future work section. The hydrolytic stability of sulfonyl fluoride inhibitors is thoroughly discussed in **Section 7.5** in the following chapter.

Lastly, confirmation of the compounds binding properties and mode of inhibition was investigated using covalent mass spectrometry experiments. As evident from the poor biochemical and cellular data, the results indicated a reversible binding mode for all synthesised compounds, ruling out covalent bond formation with the targeted catalytic Lys745.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Table 6.2: Biochemical and cellular evaluation results for the synthesised final compounds.

Compound	EGFR HTRF IC ₅₀ [nM]			EGFR CTG EC ₅₀ [nM]				
	wild-type	L858R	L858R/ T790M	A431	H1975	PC9/ T790M/ C797S	BaF3/ L858R	BaF3 T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	756 ± 340	16.0 ± 5.00	3109 ± 648	< 14	2041 ± 590
27	75.9 ± 12.8	15.7 ± 1.33	19.0 ± 13.1	7981 ± 3270	7653 ± 850	6880 ± 2550	-	-
41	225 ± 54.1	45.6 ± 3.12	84.7 ± 2.64	18040 ± 12800	11120 ± 4120	10450 ± 2980	-	-
53	365 ± 141	46.1 ± 8.62	38.4 ± 2.07	7373 ± 1920	7450 ± 2210	5909 ± 2510	< 14	13590 ± 5070
56	122 ± 30.1	14.6 ± 0.74	12.6 ± 0.60	16220 ± 9010	10930 ± 2420	4820 ± 1090	< 14	5706 ± 530
63	-	-	-	30000 ± 0	30000 ± 0	30000 ± 0	-	-

6.6 Conclusions

Building upon the previous chapter, the primary focus of our research efforts was centred on further expanding the irreversible targeting of the EGFR catalytic lysine residue and the feasibility thereof. To undertake this, we envisaged use of an osimertinib derived driving group scaffold and incorporation of various electrophiles that were deemed more reactive than the previously used 1,4-dicarbonyls. Following an arduous synthetic investigation, which required numerous strategies and various protecting groups, we established an optimised route to both the propionamide and acrylamide containing driving groups **27** and **40** respectively. Both compounds were obtained in high overall yield, allowing for easy access in their potential in the future.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Furthermore, during development of the optimised synthetic pathway, we were able to isolate several interesting intermediates of which we submitted four for biochemical and cellular evaluation.

For this initial study, we elected to make use of scaffold **27** exclusively. However, incorporation of the electrophile within the driving group was plagued by inertness. This was caused by the unexpectedly poor nucleophilicity of the indole nitrogen centre, making the overall synthesis of our final compounds extremely challenging. The consequently harsh reaction conditions required for derivatisation, such as strong base and higher temperatures, led to encounters and further compounding of other difficulties. These arose from the highly reactive nature of the electrophilic fragments and undesirable interactions from the scaffold solubilising group which will be elaborated on further in the following chapter. However, despite these complications, we were able to synthesise four final compounds containing Michael acceptor, nitrile and sulfonyl fluoride moieties. Moreover, improvements of methodology and new procedures to synthesise the unafforded compounds can be found in the future work section.

Evaluation of the synthesised precursor compounds revealed unremarkable activity for the propionamide derived intermediates. However, the Michael acceptor-containing Boc-protected analogue **39** displayed incredible potency, with values superior osimertinib. This was hypothesised to occur via metabolic cleavage of the Boc-protecting to yield the documented potent inhibitor **40**, which we also synthesised but did not submit for evaluation. While we speculate that this unintentional discovery could hold promise in the development of inhibitors containing readily metabolizable protecting groups, it is out of the scope of this project and will therefore be disregarded.

Unfortunately, the synthesised final compounds were not able to covalently modify the catalytic Lys795 residue, as indicated by covalent mass spectrometry experiments. However, the best performing nitrile-bearing compound **56** did show improved activity over the model compound **27** biochemically, and against the clinically relevant triple mutant in a cellular context, warranting its further investigation in targeting EGFR Lys745. Being a rudimentary study, we believe there is potential in pursuing targeted irreversible inhibition of this residue. However, the initial envisaged compounds must first be successfully synthesised and assessed both biochemically and in a cellular setting. This will give an indication on the viability of this strategy and whether the exploration and synthesis of additional lysine-targeting electrophiles, on this and other scaffold systems, is justifiable. Should these compounds prove ineffective or unable to undergo covalent bond formation, we propose a model reactivity study of these electrophiles be undertaken. While unable to mimic biological conditions, this study would entail reaction of the warhead bearing final compounds with a nucleophilic amine such as *p*-nitroaniline. As mentioned in the previous chapter (**Section 5.2.2**), this approach was successfully used by Kornienko and co-workers in their chemical feasibility study of polygodial.³⁸ Subsequent examination of product formation, through either isolation or analysis conducted by mass spectrometry, would provide valuable information regarding the reactivity of these lysine-targeting electrophiles and their potential for covalent bond formation.

The work undertaken in this chapter was carried out concurrently with that of the following chapter and was accomplished at the end of our entire research campaign. As such, time constraints played a large role in our decision making. With the uncovering of the synthetic complications encountered throughout this chapter, we decided to commit more time and effort towards the research endeavours of the next chapter.

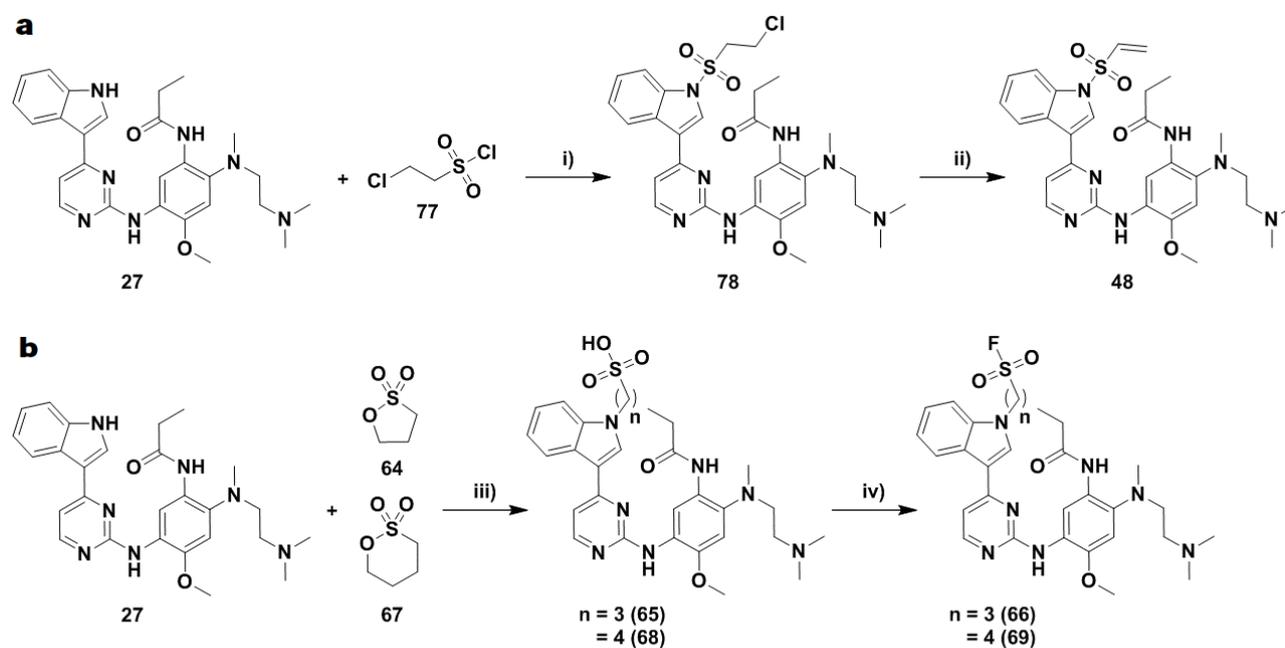
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

The successes of the following chapter, which build upon targeting unconventional residues using novel electrophiles, were made largely possible by the synthetic achievements of this chapter. After discussion of the future efforts to be undertaken for this body of research, we will delve directly into examination of the ensuing chapter.

6.7 Future Work

In light of the difficulties faced during the synthesis of our final compounds in this chapter, we have compiled various improvements to the synthetic procedures employed to potentially furnish these unsynthesised inhibitors. These refinements may be found in **Scheme 6.14**. To gain access to the ethenesulfonamide-containing compound **48**, we ordered the commercially available chloro-variant electrophile **77**, as shown in **Scheme 6.14a**. As alluded to in **Section 6.4.1.2**, with the passing of new legislation at the end of 2017, purchasing and shipping of reactive and flammable reagents to South Africa has skyrocketed the costs involved and lengthened transportation time to periods exceeding six-months. For this reason, we were unable to originally synthesise **48**, as we were waiting for delivery of this reagent and hesitated to partake in the hazardous synthesis of this building block.

With this electrophile in hand, we plan to sulfonate the indole nitrogen of scaffold **27** via deprotonation with sodium hydride at very low temperatures. It is our hope that substitution of the bromo- for a chloro-leaving group will diminish the reactivity of the system enough to prevent spontaneous elimination which is believed to have occurred in our attempted synthesis. Following isolation of intermediate **78**, which could also be considered for evaluation in targeting the Lys745, we would promote elimination through treatment with excess triethylamine. Use of this procedure is well-documented in the literature and could conceivably afford the sought-after inhibitor **48**.



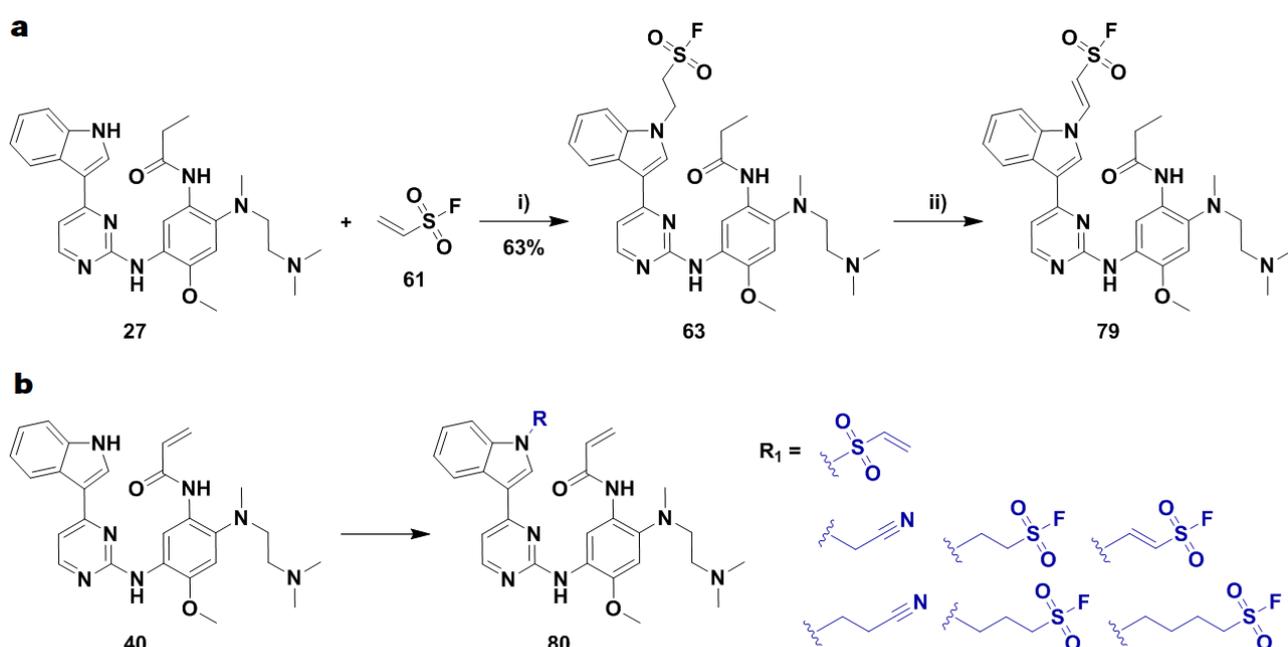
Scheme 6.14: Proposed synthetic alteration to potentially afford a) ethenesulfonamide **48** and b) propyl and butyl bridged sulfonyl fluorides **66** and **69** respectively. Reagents and conditions: i) NaH (1.05 equiv.), DMF, $-40\text{ }^{\circ}\text{C}$ – rt, 30 min then **77** (1.2 equiv.), DMF, $-40\text{ }^{\circ}\text{C}$ – $-20\text{ }^{\circ}\text{C}$, 2 h; ii) Et₃N (2 equiv.), DCM, $-10\text{ }^{\circ}\text{C}$ – rt, 2 h; iii) NaH (1.05 equiv.), DMF, $-40\text{ }^{\circ}\text{C}$ – rt, 30 min then **64** or **67** (1.5 equiv.), DMF, $-10\text{ }^{\circ}\text{C}$ – rt, 2 h; iv) DAST (1.05 equiv.), DCM, $-78\text{ }^{\circ}\text{C}$, 1h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

As encountered in **Section 6.4.2.3**, all efforts to synthesise and isolate the propyl and butyl bridged sulfonyl fluorides, **66** and **69** respectively in **Scheme 6.14b**, were met with failure. This was mainly ascribed to impracticalities surrounding purification of the water-soluble intermediates **65** and **68**. Consulting with colleagues at the biochemistry department at Stellenbosch University, we were pointed towards the use of Amberlyst 15 resin.

This resin acts as a sulfonic ion-exchange acid resin and may serve as a plausible option for purification of these sulfonate derived compounds. Furthermore, should this method prove unsuccessful, we intend to utilise reverse-phase column chromatography (C18 bonded silica gel) as an alternate plan. Therefore, application of the original reaction conditions that were suspected to have worked will afford the crude sulfonate reaction mixture. Implementation of either of the above discussed methods of purification could then furnish pure samples of **65** and **68**. These compounds will then be exposed to the fluorinating reagent DAST at low temperatures, potentially affording the desired final compounds **66** and **69**. We believe that synthesis of these particular compounds are imperative for the success of this research venture, as the only known covalent modification of the EGFR catalytic lysine was accomplished through a sulfonyl fluoride warhead.¹

Another priority is the resynthesis and submission for screening of final compound **63**, suspected to have degraded/hydrolysed en route to our collaborators in Germany. The synthesis of **63** will be carried out according to the parameters shown in **Scheme 6.15a**, where after it will be carefully stored and delivered by express courier. Additionally, we propose synthesis of the unsaturated sulfonyl fluoride compound **79**, another electrophile capable of potentially undergoing covalent bond formation with the catalytic Lys745 residue. In the following chapter, use of excess manganese dioxide to facilitate oxidation of the activated ethylene bridge to the alkene was successfully carried out on a structurally related analogue.



Scheme 6.15: a) Resynthesis of the presumed degraded final compound **63** and proposed synthesis of unsaturated sulfonyl fluoride compound **79** and b) Michael acceptor containing final compounds targeting the EGFR-L858R/T790M double mutant.

Reagents and conditions: i) ESF (1.3 equiv.), K₂CO₃ (1.5 equiv.), DMF, 0 °C – rt, 6 h; ii) MnO₂ (15 equiv.), CHCl₃, rt, 8 h.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

This generated a “dual warhead” or compound containing two electrophilic centres, which forms a major part of our therapeutic targeting strategy and will be comprehensively discussed as **Section 7.4.3.2** in the next chapter. Therefore, we wish to reproduce this synthetic approach to potentially yield the dual warhead incorporated inhibitor **79** for irreversible targeting of the catalytic lysine residue.

Following the effective synthesis of all originally planned inhibitors derived from the propionamide driving group **27**, we seek to replicate the synthetic methodology utilising the Michael acceptor-containing heterocyclic scaffold **40**. Illustrated in **Scheme 6.15b**, this will furnish a similar group of final compounds (**80**) with the enhanced ability of targeting the Cys797 residue covalently. These compounds can then be evaluated explicitly against the double L858R/T790M mutant variant, potentially improving on the inhibitory properties of osimertinib through supplementary interactions at the roof of the enzyme. Furthermore, covalent bond formation will provide enhanced cellular binding and longer residence, increasing the likelihood of our incorporated lysine-targeting electrophiles to covalently modify Lys745. Regardless of the outcome, evaluation of these inhibitors will provide valuable information toward the development of this research endeavour.

Though not visually represented, we chose to abandon any further pursuit in the synthesis of α - and β -haloketone containing compounds. This decision is based on the complications surrounding their production, both in this chapter and in the following, which is mainly ascribed to their highly reactive nature. In a recent study, these functional groups were shown to be exceedingly promiscuous, leading to off-target protein modification.³⁹ Therefore, we are not in a position to endorse and have not included any synthesis of these electrophiles in our future considerations.

Lastly, we looked towards expansion of this research thrust to two other heterocyclic driving scaffolds. The first of these, displayed in **Figure 6.10** on the following page, draws inspiration from the dual ALK/EGFR inhibitor **11**. As outlined in **Section 6.2.2**, molecular modelling of compound **11** revealed the isopropylsulfone moiety to be in a proximal location to the catalytic Lys745 residue, resulting in reversible interactions with the hydrogen bond-accepting carbonyl groups. This compound bears a striking resemblance to the osimertinib scaffold, with the exception of a substituted phenyl ring in place of the indole group and a chloro-substituent on the 5-position of the pyrimidine heterocycle. This similarity will allow for easy synthetic access to the envisaged group of final compounds (**81**), containing the various electrophiles (R_1 , blue) capable of labelling the catalytic lysine residue.

In lieu of the electrophiles being attached to a more nucleophilic aniline fragment rather than the unreactive indole nitrogen, the final installation step should be far less complicated. Furthermore, this would facilitate inclusion of the cyanoacetamide moiety and allow for use of a sulfonyl fluoride directly attached to the phenyl ring. Once again, we would synthesise the propionamide and Michael acceptor analogues (R_2 , red), with the option of substituting the troublesome osimertinib solubilising group with the widely used alternative N-methylpiperazine moiety (R_3 , green). The combination of these features makes this scaffold a prime candidate for developing a library of lysine-targeting compounds.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

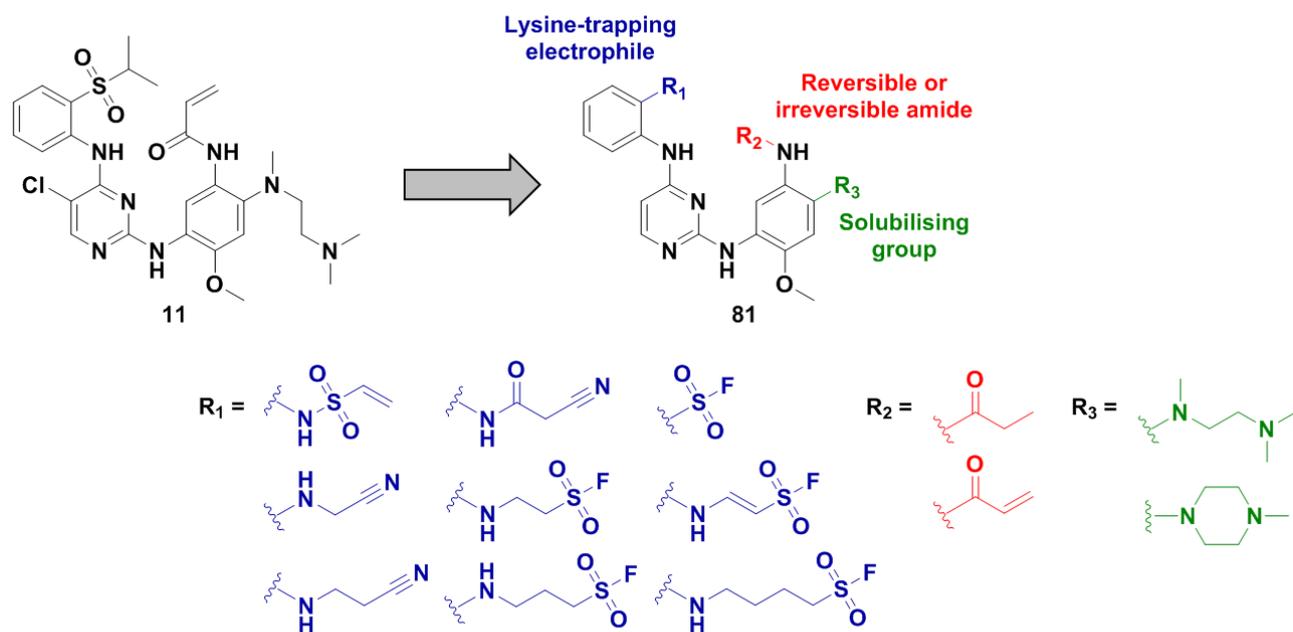


Figure 6.10: Proposed library of lysine-targeting compounds inspired by the dual ALK/EGFR inhibitor **11**.

The final consideration and second driving group that could be employed will be based on the gefitinib core utilised in Chapter 4. As summarised in **Section 5.7**, the synthesis of heterocyclic scaffold **82** (**Figure 6.11**), which contains the gefitinib solubilising group, will accommodate derivatisation at the 4-amino position of the quinazoline ring. This will allow for either direct substitution of electrophiles (R_1 , blue) on the amino quinazoline (**83**) or incorporation at the 3- or 4-position (R_2 , blue) of an installed aniline ring (**84**). Synthesis of these compounds, with warheads installed at various positions, will allocate a wide range of exploitable chemical space for the electrophiles, thereby increasing the probability for covalent modification of the EGFR catalytic lysine residue. This strategy can be used in conjunction with the incorporation of 1,4-dicarbonyl electrophiles laid out in Chapter 5.

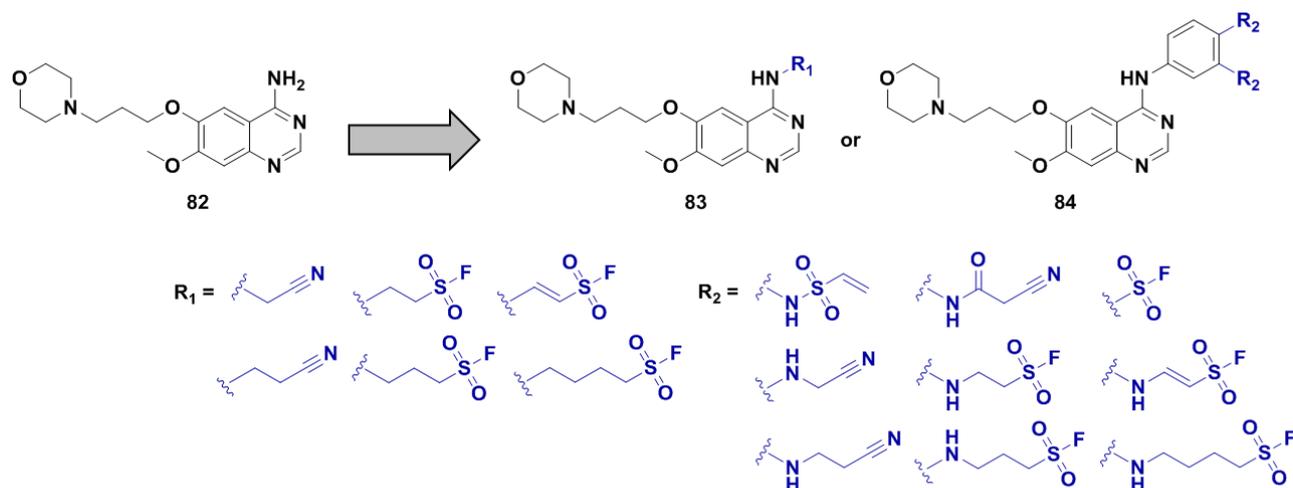
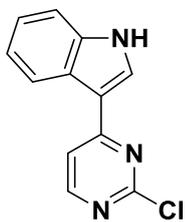


Figure 6.11: Proposed library of lysine-targeting compounds building on the gefitinib scaffold **82**.

6.8 Supplementary Information

3-(2-Chloropyrimidin-4-yl)-1H-indole (14)²⁴

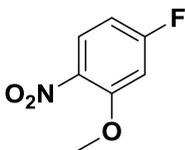


A 1 M solution of ethylmagnesium bromide was prepared as follows: A flame-dried, 2-neck round-bottomed flask was charged with oven-dried magnesium turnings (1.54 g, 64.0 mmol, 2 equiv.) and purged with Ar. To this was added THF (48 mL), followed by the dropwise addition of previously distilled bromoethane (5.23 g, 48.0 mmol, 1.5 equiv.) which resulted in an exotherm. The solution was then allowed to stir for 45 min, during which time the solution cooled and developed a grey colour. A separate 2-neck round-bottom flask was charged with indole (5.62 g, 48.0 mmol, 1.5 equiv.) and purged with Ar. To this was added THF (50 mL) and the solution was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice). The prepared 1 M ethylmagnesium bromide solution (48.0 mL, 48.0 mmol, 1.5 equiv.) was then added dropwise to this solution over 20 minutes, being careful to keep the temperature below $0\text{ }^{\circ}\text{C}$. The solution was then allowed to warm to rt over 90 minutes, resulting in the formation of a white precipitate. The suspension was then recooled to $-10\text{ }^{\circ}\text{C}$ and 2,4-dichloropyrimidine (4.78 g, 32.0 mmol, 1 equiv.) was added as a single portion with no exotherm. The reaction mixture was placed at $80\text{ }^{\circ}\text{C}$ and allowed to stir for 5 h, undergoing a colour change from orange to deep red. After allowing to cool to rt, the reaction was quenched by the slow addition of MeOH (2.00 mL, 48.0 mmol, 1.5 equiv.) resulting in an exotherm. After removal of the solvent *in vacuo*, the crude product was adsorbed onto silica and purification was achieved using flash column chromatography with elution gradient of 0 – 5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **14** as a white-yellow powder (4.48 g, 19.5 mmol, 61%).

¹H NMR and MS data collected for this compound compared well with the reported literature values.²⁴

Rf: 0.51 (1 % MeOH in DCM); ¹H NMR (300 MHz, DMSO-*d*₆) δ 7.25 – 7.21 (m, 2H, 2 \times ArH), 7.53 – 7.49 (m, 1H, ArH), 7.90 (d, *J* = 5.5 Hz, 1H, ArH), 8.40 – 8.44 (m, 1H, ArH), 8.50 (s, 1H, ArH), 8.52 (d, *J* = 5.5 Hz, 1H, ArH), 12.08 (br s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 111.9, 112.36, 114.5, 121.4, 121.6, 122.7, 124.9, 131.1, 137.3, 158.5, 160.3, 165.0 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₂H₉N₃Cl: 230.0485; found: 230.0480.

4-Fluoro-2-methoxy-1-nitrobenzene (19)²⁴

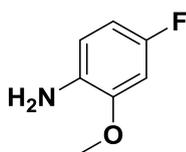


A 2-neck round-bottomed flask was charged with 5-fluoro-2-nitrophenol (8.00 g, 50.9 mmol, 1 equiv.) and potassium carbonate (8.45 g, 61.1 mmol, 1.2 equiv.) and purged with Ar. To this was added acetone (150 mL) and the resulting suspension was allowed to stir for 30 min at rt. The suspension was cooled to $0\text{ }^{\circ}\text{C}$ and methyl iodide (3.49 mL, 56.0 mmol, 1.1 equiv.) was added dropwise. The reaction mixture was placed at $60\text{ }^{\circ}\text{C}$ and allowed to stir for 12 h, during which time a deep red colour developed. After allowing to cool to rt, the solvent was removed *in vacuo* and the reaction mixture was diluted with EtOAc (50 mL) and H₂O (100 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 \times 50 mL) and the combined organic layers were washed with a 1 M solution of NaOH (2 \times 50 mL), a saturated solution of brine (50 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, compound **19** was afforded as a beige solid (8.54 g, 49.9 mmol, 98%) with no further purification required.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

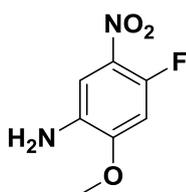
^1H NMR data collected for this compound compared well with the reported literature value.²⁴

Rf: 0.66 (30 % EtOAc in PE); ^1H NMR (600 MHz, CDCl_3) δ 3.96 (s, 3H, ArOCH_3), 6.74 – 6.70 (m, 1H, ArH), 6.79 (dd, $J = 10.8, 2.7$ Hz, 1H, ArH), 7.95 (dd, $J = 8.6, 6.1$ Hz, 1H, ArH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 57.0, 101.5 (d, $J = 27.1$ Hz), 107.4 (d, $J = 23.5$ Hz), 128.3 (d, $J = 11.5$ Hz), 155.5 (d, $J = 11.4$ Hz), 165.1, 166.8 ppm; no ionization was found in MS as reported in the literature.

4-Fluoro-2-methoxyaniline (20)

A 3-neck round-bottom flask was charged with 4-fluoro-2-methoxy-1-nitrobenzene (6.00 g, 35.1 mmol, 1 equiv.) and Pd/C (375 mg, 3.51 mmol, 10 mol%), placed under vacuum and EtOH (35 mL) was added. A hydrogen atmosphere was introduced via balloon and the suspension was then allowed to stir at rt for 24 h or until complete consumption of starting material as indicated by TLC. The reaction mixture was filtered through a plug of celite, washed with EtOH (35 mL) and the solvent was removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 10 – 30% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **20** as a brown oil (4.55 g, 32.3 mmol, 92%).

Rf: 0.52 (30 % EtOAc in PE); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 3.76 (s, 3H, ArOCH_3), 4.53 (br s, 2H, ArNH_2), 6.49 (tdd, $J = 8.6, 2.7, 1.3$ Hz, 1H, ArH), 6.59 (ddd, $J = 8.6, 6.1, 1.3$ Hz, 1H, ArH), 6.71 (dt, $J = 10.8, 1.3$ Hz, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 55.6, 99.1 (d, $J = 26.7$ Hz), 105.9 (d, $J = 21.4$ Hz), 113.2 (d, $J = 9.1$ Hz), 133.9 (d, $J = 2.2$ Hz), 146.8 (d, $J = 9.7$ Hz), 152.9; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_7\text{H}_9\text{NOF}$: 142.0668; found: 142.0668.

4-Fluoro-2-methoxy-5-nitroaniline (21)

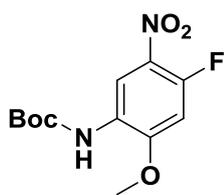
To a solution of 85% aqueous H_2SO_4 (v/v 35.0 mL, 531 mmol, 25 equiv.), cooled to -10 °C (acetone/ice), was added 4-fluoro-2-methoxyaniline (3.00 g, 21.3 mmol, 1 equiv.) dropwise over 10 min. The solution was then allowed to stir at -10 °C for 30 min followed by portion wise addition of KNO_3 (2.15 g, 21.3 mmol, 1 equiv.) over 15 min. The reaction mixture was then allowed to stir at 0 °C for 4 h and poured onto a mixture of ice/water (70 mL). The slurry was

carefully neutralized with concentrated NH_4OH and the resulting precipitate was collected by filtration, washed with water (3×20 mL), PE (2×20 mL) and dried *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 30 – 90% DCM in PE. Pure fractions were evaporated to dryness to afford compound **21** as an orange powder (3.48 g, 18.7 mmol, 88%).

^1H NMR and MS data collected for this compound compared well with the reported literature values.²⁵

Rf: 0.22 (20 % DCM in PE); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 3.90 (s, 3H, ArOCH_3), 5.22 (br s, 2H, ArNH_2), 7.02 (d, $J = 13.4$ Hz, 1H, ArH), 7.34 (d, $J = 8.0$ Hz, 1H, ArH) ppm; ^{13}C NMR (75 MHz, DMSO) δ 56.7, 100.4 (d, $J = 26.5$ Hz), 106.9 (d, $J = 2.2$ Hz), 134.8 (d, $J = 1.8$ Hz), 146.6, 150.0, 152.1 (d, $J = 9.0$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_7\text{H}_8\text{N}_2\text{O}_3\text{F}$: 187.0519; found: 187.0524.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

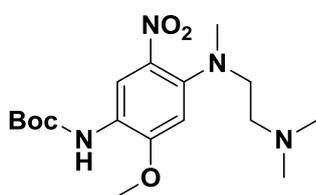
tert-Butyl (4-fluoro-2-methoxy-5-nitrophenyl)carbamate (22)⁴⁰

A 2-neck round-bottom flask was charged with 4-fluoro-2-methoxy-5-nitroaniline (3.00 g, 16.1 mmol, 1 equiv.), purged with Ar, dissolved in DCM (35 mL) and cooled to 0 °C. A solution of Boc₂O (3.69 g, 16.9 mmol, 1.05 equiv.) in DCM (16 mL) was added dropwise, followed by the addition of a catalytic amount of DMAP. The reaction mixture was then

allowed to warm to rt and stirred for 24 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 30 – 80% DCM in PE. Pure fractions were evaporated to dryness to afford compound **22** as a white-yellow solid (3.78 g, 13.2 mmol, 82%).

Mp, ¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.⁴⁰

Rf: 0.61 (50 % DCM in PE); mp: 88 – 90 °C; ¹H NMR (300 MHz, DMSO-*d*₆) δ 1.47 (s, 9H, OC(CH₃)₃), 3.94 (s, 3H, ArOCH₃), 7.27 (d, *J* = 13.3 Hz, 1H, ArH), 8.50 (s, 1H, ArNHBoc), 8.52 (d, *J* = 8.0 Hz, 1H, ArH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 27.9, 57.2, 80.0, 101.3 (d, *J* = 26.2 Hz), 116.6, 124.4 (d, *J* = 2.7 Hz), 128.5 (d, *J* = 6.6 Hz), 150.8, 152.7, 154.2, 155.6 (d, *J* = 10.2 Hz) ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₂H₁₄N₂O₅F: 285.0887; found: 285.0891.

tert-Butyl (4-fluoro-2-methoxy-5-nitrophenyl)carbamate (23)⁴⁰

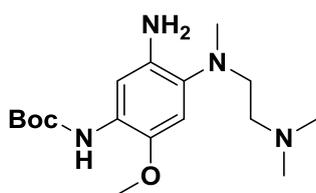
A 2-neck round-bottom flask was charged with *tert*-butyl (4-fluoro-2-methoxy-5-nitrophenyl)carbamate (3.38 g, 11.8 mmol, 1 equiv.), purged with Ar and dissolved in DMA (30 mL). To this solution was added DIPEA (2.16 mL, 12.4 mmol, 1.05 equiv.), followed by the slow addition of a solution of *N*¹,*N*¹,*N*²-trimethylethane-1,2-diamine (1.51 g, 14.8 mmol, 1.25 equiv.) in DMA (5 mL) at rt. The reaction was

placed at 60 °C and allowed to stir for 4 h, during which time the reaction mixture developed a deep red colour. After allowing to cool to rt, the reaction mixture was then diluted with H₂O (70 mL) and EtOAc (70 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 × 30 mL) and the combined organic layers were then washed with successive aliquots of a saturated solution of brine (3 × 30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **23** as an orange powder (4.31 g, 11.7 mmol, 99%).

Mp, ¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.⁴⁰

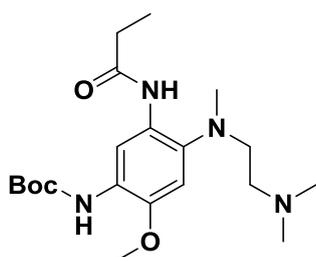
Rf: 0.43 (10 % MeOH in DCM); mp: 92 – 94 °C; ¹H NMR (300 MHz, DMSO-*d*₆) δ 1.45 (s, 9H, OC(CH₃)₃), 2.16 (s, 6H, CH₂N(CH₃)₂), 2.47 (t, *J* = 6.8 Hz, 2H, NCH₂CH₂N), 2.80 (s, 3H, ArNCH₃), 3.23 (t, *J* = 6.8 Hz, 2H, NCH₂CH₂N), 3.90 (s, 3H, ArOCH₃), 6.74 (s, 1H, ArH), 8.14 (br s, 2H, ArH and ArNHBoc) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 28.0, 40.4, 45.3, 52.5, 56.2, 79.4, 101.9, 119.1, 119.6, 131.7, 144.3, 153.0, 154.7 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₇H₂₉N₄O₅: 369.2138; found: 369.2143.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

tert-Butyl (5-amino-4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxyphenyl)carbamate (24)

A 3-neck round-bottom flask was charged with *tert*-butyl (4-fluoro-2-methoxy-5-nitrophenyl)carbamate (4.20 g, 11.4 mmol, 1 equiv.) and Pd/C (121 mg, 1.14 mmol, 10 mol%), placed under vacuum and MeOH (50 mL) was added. A hydrogen atmosphere was introduced *via* balloon and the suspension was then allowed to stir at rt for 12 h or until complete consumption of starting material as indicated by TLC. The reaction mixture was filtered through a plug of celite, washed with MeOH (35 mL) and the solvent was removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 10 – 30% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **24** as a brown oil (3.51 g, 10.4 mmol, 91%).

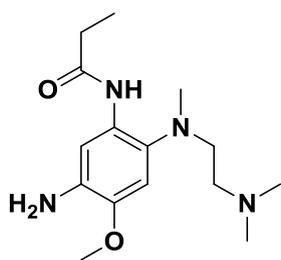
Rf: 0.19 (10 % MeOH in DCM); ¹H NMR (300 MHz, DMSO-*d*₆) δ 1.44 (s, 9H, OC(CH₃)₃), 2.51 (s, 3H, ArNCH₃), 2.60 (s, 6H, CH₂N(CH₃)₂), 3.14 – 2.97 (m, 4H, 2 × NCH₂CH₂N), 3.69 (s, 3H, ArOCH₃), 6.69 (s, 1H, ArH), 7.09 (s, 1H, ArH), 7.57 (s, 1H, ArNH₂Boc) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 28.1, 41.4, 42.7, 50.4, 54.2, 56.5, 78.8, 105.8, 108.0, 124.25, 133.0, 137.2, 141.1, 152.7 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₇H₃₁N₄O₃: 339.2396; found: 339.2401.

tert-Butyl (4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxy-5-propionamidophenyl)carbamate (25)

A 2-neck round-bottom flask was charged with *tert*-butyl (5-amino-4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxyphenyl)carbamate (1.35 g, 3.99 mmol, 1 equiv.), purged with Ar and dissolved in DCM (30 mL). The solution was cooled to –10 °C (acetone/ice) and DIPEA (0.760 mL, 4.39 mmol, 1.1 equiv.) was added. A solution of propionyl chloride (387 mg, 4.19 mmol, 1.05 equiv.) in DCM (5 mL) was then added dropwise to the solution over 10 min. The reaction mixture was then allowed to warm to rt and stirred for 2 h. The reaction mixture was then quenched and diluted with H₂O (30 mL), a saturated solution of NaHCO₃ (30 mL) and DCM (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 30 mL) and the combined organic layers were then washed with a saturated solution of brine (50 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **25** as a brown foam (1.43 g, 3.63 mmol, 89%).

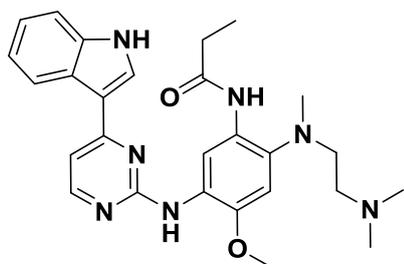
Rf: 0.37 (10 % MeOH in DCM); ¹H NMR (300 MHz, CDCl₃) δ 1.24 – 1.12 (m, 3H, COCH₂CH₃), 1.47 (s, 9H, OC(CH₃)₃), 2.25 (s, 6H, CH₂N(CH₃)₂), 2.38 – 2.27 (m, 4H, NCH₂CH₂N and COCH₂CH₃), 2.61 (s, 3H, ArNCH₃), 2.84 (br s, 2H, NCH₂CH₂N), 3.77 (s, 3H, ArOCH₃), 6.65 (s, 1H, ArH), 6.86 (s, 1H, ArH), 8.92 (s, 1H, ArNH₂Boc), 9.63 (s, 1H, ArNHCO) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 10.0, 28.4, 30.5, 43.8, 45.2, 55.8, 56.0, 57.1, 80.3, 104.4, 111.5, 125.3, 129.6, 136.0, 144.1, 152.6, 171.9 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₀H₃₅N₄O₄: 395.2658; found: 395.2665.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

N-(5-Amino-2-[[2-(dimethylamino)ethyl](methyl)amino]-4-methoxyphenyl]propionamide (26)

A 2-neck round-bottom flask was charged with *tert*-butyl (4-[[2-(dimethylamino)ethyl](methyl)amino]-2-methoxy-5-propionamidophenyl)carbamate (1.22 g, 3.09 mmol, 1 equiv.), purged with Ar and dissolved in DCM (15 mL). The solution was cooled to 0 °C and trifluoroacetic acid (3 mL) was added dropwise to the solution. The reaction mixture was then allowed to warm to rt and stirred for 6 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, the residue was diluted with DCM (50 mL) and a 1 M solution of NaOH (50 mL) and the organic layer was separated. The organic layer was then washed with a saturated solution of NaHCO₃ (30 mL), a saturated solution of brine (50 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **26** as a brown oil (782 mg, 2.66 mmol, 86%).

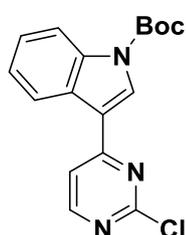
Rf: 0.25 (10 % MeOH in DCM); ¹H NMR (300 MHz, CDCl₃) δ 1.18 (t, *J* = 7.6 Hz, 3H, COCH₂CH₃), 2.20 – 2.15 (m, 8H, CH₂N(CH₃)₂ and NCH₂CH₂N), 2.30 (q, *J* = 7.6 Hz, 2H, COCH₂CH₃), 2.58 (s, 3H, ArNCH₃), 2.79 (t, *J* = 6.0 Hz, 2H, NCH₂CH₂N), 3.71 (br s, 2H, ArNH₂), 3.74 (s, 3H, ArOCH₃), 6.61 (s, 1H, ArH), 7.81 (s, 1H, ArH), 9.67 (s, 1H, ArNHCO) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 10.0, 30.6, 44.2, 45.4, 55.8, 56.3, 57.3, 105.1, 106.8, 130.1, 131.8, 133.8, 143.2, 171.9 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₅H₂₇N₄O₂: 295.2134; found: 295.2140.

N-(5-{[4-(1*H*-Indol-3-yl)pyrimidin-2-yl]amino}-2-[[2-(dimethylamino)ethyl](methyl)amino]-4-methoxyphenyl]propionamide (27)

A 2-neck round-bottomed flask was charged with 3-(2-chloropyrimidin-4-yl)-1*H*-indole (96.0 mg, 0.42 mmol, 1 equiv.) and *N*-(5-amino-2-[[2-(dimethylamino)ethyl](methyl)amino]-4-methoxyphenyl)propionamide (123 mg, 0.418 mmol, 1 equiv.) and purged with Ar. To this was added 2-pentanol (10 mL) and the resulting suspension was placed at 105 °C. *p*-Toluenesulfonic acid monohydrate (96.0 mg, 0.505 mmol, 1.2 equiv.) was added as a single portion and the reaction mixture was then allowed to stir for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then allowed to cool and stand for 8 h which resulted in the formation of a brown precipitate. The precipitate was collected by filtration, washed with aliquots of acetonitrile (3 × 30 mL) and dried *in vacuo*. The solid was dissolved in 10% MeOH in DCM and washed with a saturated solution of NaHCO₃ (3 × 30 mL), a saturated solution of brine (30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **27** as a beige solid (126 mg, 0.258 mmol, 61%).

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

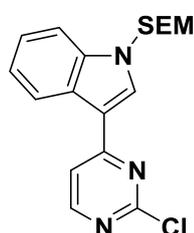
Rf: 0.21 (10 % MeOH in DCM); IR (ATR) cm^{-1} : 3412 (m, N-H stretch), 3171 (m, N-H stretch), 2953 (m, C-H stretch), 2819 (m, C-H stretch), 1656 (s, C=O stretch), 1577 (m, N-H bend), 1447 (m, C-H bend), 1400 (m, C-H bend), 1340 (s, C-N stretch), 1056 (s, C-O stretch), 741 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 1.30 (t, $J = 7.6$ Hz, 3H, COCH_2CH_3), 2.31 – 2.21 (m, 8H, $\text{CH}_2\text{N}(\text{CH}_3)_2$ and $\text{NCH}_2\text{CH}_2\text{N}$), 2.45 (q, $J = 7.6$ Hz, 2H, COCH_2CH_3), 2.67 (s, 3H, ArNCH_3), 2.88 (t, $J = 5.8$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.85 (s, 3H, ArOCH_3), 6.76 (s, 1H, ArH), 6.96 (d, $J = 5.3$ Hz, 1H, ArH), 7.18 – 7.10 (m, 2H, $2 \times \text{ArH}$), 7.32 (d, $J = 7.2$ Hz, 1H, ArH), 7.65 (s, 1H, ArH), 8.22 – 8.15 (m, 2H, $2 \times \text{ArH}$), 8.37 (s, 1H, ArH), 9.64 (s, 1H, ArNH), 9.97 (s, 1H, ArNH), 10.38 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 10.3, 30.8, 44.1, 45.3, 56.0, 56.1, 57.3, 104.7, 108.2, 111.0, 112.1, 114.7, 120.8, 120.8, 122.0, 125.4, 127.5, 129.4, 129.7, 134.9, 137.5, 144.6, 157.3, 159.8, 162.8, 172.3 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{34}\text{N}_7\text{O}_2$: 488.2774; found: 488.2783.

tert-Butyl 3-(2-chloropyrimidin-4-yl)-1H-indole-1-carboxylate (28)

A 2-neck round-bottom flask was charged with 3-(2-chloropyrimidin-4-yl)-1H-indole (500 mg, 2.18 mmol, 1 equiv.), purged with Ar, dissolved in DCM (20 mL) and Et_3N (0.390 mL, 2.83 mmol, 1.3 equiv.) was added. After cooling to 0°C , a solution of Boc_2O (570 mg, 2.61 mmol, 1.2 equiv.) in DCM (10 mL) was added dropwise, followed by the addition of a catalytic amount of DMAP.

The reaction mixture was then allowed to warm to rt and stirred for 12 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 50 – 90% DCM in PE. Pure fractions were evaporated to dryness to afford compound **28** as a yellow solid (639 mg, 1.94 mmol, 89%).

Rf: 0.69 (DCM); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 1.67 (s, 9H, $\text{OC}(\text{CH}_3)_3$), 7.45 – 7.34 (m, 2H, $2 \times \text{ArH}$), 8.11 (dd, $J = 7.0, 2.0$ Hz, 1H, ArH), 8.14 (d, $J = 5.4$ Hz, 1H, ArH), 8.47 (dd, $J = 7.0, 2.0$ Hz, 1H, ArH), 8.68 (dd [app. t], $J = 2.7$ Hz, 2H, $2 \times \text{ArH}$) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 27.6, 85.2, 114.9, 116.2, 116.3, 122.2, 123.8, 125.3, 126.7, 130.0, 135.4, 148.5, 159.6, 160.1, 163.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{17}\text{H}_{17}\text{N}_3\text{O}_2\text{Cl}$: 330.1009; found: 330.0999.

3-(2-Chloropyrimidin-4-yl)-1-{[2-(trimethylsilyl)ethoxy]methyl}-1H-indole (30)

A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in mineral oil, 465 mg, 11.6 mmol, 1.16 equiv.), purged with Ar and to this was added DMF (30 mL). The resulting suspension was cooled to -10°C (acetone/ice) and a solution of 3-(2-chloropyrimidin-4-yl)-1H-indole (2.30 g, 10.0 mmol, 1 equiv.) in DMF (10 mL) was added dropwise over 15 min. The solution was then allowed to warm to rt and stirred for 30 min. After

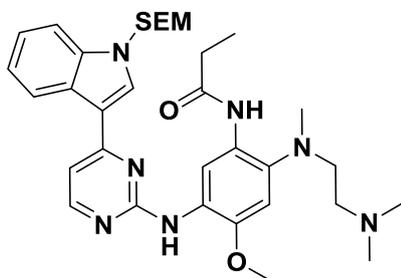
recooling the solution to -10°C (acetone/ice), a solution of SEM-Cl (2.05 g, 12.3 mmol, 1.23 equiv.) in DMF (10 mL) was added dropwise over 10 min. The reaction mixture was then allowed to warm to rt and stirred for 4 h or until complete consumption of the starting material as indicated by TLC, changing from yellow to colourless during this period. The reaction mixture was then quenched and diluted with H_2O (50 mL) and EtOAc (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×30 mL) and the combined organic layers were then washed with a saturated solution of brine (3×50 mL), dried over MgSO_4 and filtered.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of DCM to 2% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **30** as an off white solid (2.74 g, 7.61 mmol, 76%).

Rf: 0.43 (DCM); ^1H NMR (300 MHz, CDCl_3) δ -0.06 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.90 (t, $J = 8.0$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 3.51 (t, $J = 8.0$ Hz, 2H, OCH_2CH_2), 5.51 (s, 2H, NCH_2O), 7.37 – 7.31 (m, 2H, *ArH*), 7.49 (d, $J = 5.4$ Hz, 1H, *ArH*), 7.55 (dd, $J = 6.4, 2.7$ Hz, 1H, *ArH*), 8.00 (s, 1H, *ArH*), 8.33 (dd, $J = 6.4, 2.7$ Hz, 1H, *ArH*), 8.45 (d, $J = 5.4$ Hz, 1H, *ArH*) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ -1.3, 17.8, 66.6, 76.4, 111.0, 113.6, 114.4, 121.6, 122.6, 123.7, 126.1, 131.6, 137.6, 158.6, 161.5, 164.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{18}\text{H}_{22}\text{N}_3\text{OSiCl}$: 360.1220; found: 360.1296.

***N*-(2-([2-(dimethylamino)ethyl](methyl)amino)-4-methoxy-5-{[4-(1-{[2-(trimethylsilyl)ethoxy]methyl}]-1*H*-indol-3-yl)pyrimidin-2-yl]amino}phenyl)propionamide (31)**

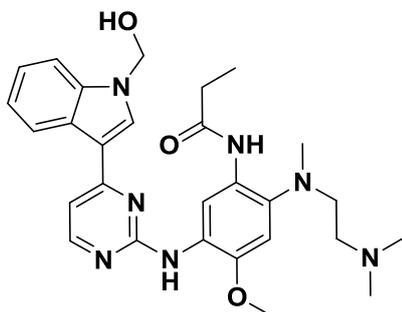


A 2-neck round-bottomed flask was charged with 3-(2-chloropyrimidin-4-yl)-1-([2-(trimethylsilyl)ethoxy]methyl)-1*H*-indole (230 mg, 0.640 mmol, 1 equiv.) and *N*-(5-amino-2-((2-(dimethylamino)ethyl)(methyl)amino)-4-methoxyphenyl)propionamide (186 mg, 0.640 mmol, 1 equiv.) and purged with Ar. To this was added 2-pentanol (10 mL) and the resulting suspension was placed at 105 °C. *p*-Toluenesulfonic acid monohydrate (145 mg, 0.762

mmol, 1.2 equiv.) was added as a single portion and the reaction mixture was then allowed to stir for 4 h or until complete consumption of the starting material as indicated by TLC. After allowing to cool to rt, the reaction mixture was diluted with DCM (30 mL), H_2O (20 mL) and neutralised with a saturated solution of NaHCO_3 (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5% – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **31** as a tan foam (302 mg, 0.489 mmol, 77%).

Rf: 0.6 10 % MeOH in DCM; IR (ATR) cm^{-1} : 3419 (m, N-H stretch), 3118 (m, N-H stretch), 2947 (m, C-H stretch), 2817 (m, C-H stretch), 1670 (s, C=O stretch), 1577 (m, N-H bend), 1454 (m, C-H bend), 1400 (m, C-H bend), 1347 (s, C-N stretch), 1069 (s, C-O stretch), 737 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ -0.08 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.90 (t, $J = 8.0$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 1.31 (t, $J = 7.6$ Hz, 3H, COCH_2CH_3), 2.30 – 2.24 (m, 8H, $\text{CH}_2\text{N}(\text{CH}_3)_2$ and $\text{NCH}_2\text{CH}_2\text{N}$), 2.45 (q, $J = 7.6$ Hz, 2H, COCH_2CH_3), 2.68 (s, 3H, ArNCH_3), 2.90 (t, $J = 5.8$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.59 (t, $J = 8.0$ Hz 2H, OCH_2CH_2), 3.87 (s, 3H, ArOCH_3), 5.79 (s, 2H, NCH_2O), 6.78 (s, 1H, *ArH*), 7.23 (d, $J = 5.3$ Hz, 1H, *ArH*), 7.30 – 7.27 (m, 2H, $2 \times \text{ArH}$), 7.63 (dd, $J = 6.3, 2.7$ Hz, 1H, *ArH*), 7.75 (s, 1H, *ArH*), 8.08 (dd, $J = 6.3, 2.7$ Hz, 1H, *ArH*), 8.40 (d, $J = 5.3$ Hz, 1H, *ArH*), 9.20 (s, 1H, *ArH*), 9.73 (s, 1H, *ArNH*), 9.75 (s, 1H, *ArNH*) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ -1.3, 10.3, 17.9, 30.9, 44.1, 45.6, 56.2, 57.5, 65.9, 76.5, 104.7, 108.2, 109.5, 111.5, 114.6, 120.4, 121.6, 122.3, 126.7, 127.7, 129.9, 134.2, 134.4, 137.68, 143.9, 158.1, 159.7, 162.1, 171.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{33}\text{H}_{48}\text{N}_7\text{O}_3\text{Si}$: 618.3588; found: 618.3601.

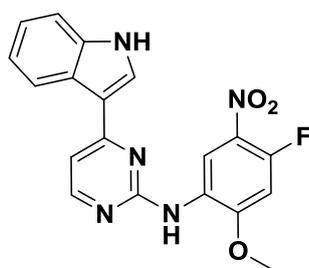
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

N-(2-([2-(dimethylamino)ethyl](methyl)amino)-5-([4-[1-(hydroxymethyl)-1H-indol-3-yl]pyrimidin-2-yl]amino)-4-methoxyphenyl)propionamide (32)

A 2-neck round-bottomed flask was charged with *N*-(2-([2-(dimethylamino)ethyl](methyl)amino)-4-methoxy-5-([4-(1-([2-(trimethylsilyl)ethoxy)methyl]-1*H*-indol-3-yl)pyrimidin-2-yl]amino)phenyl)propionamide (140 mg, 0.227 mmol, 1 equiv.), purged with Ar and dissolved in DCM (9 mL). This was followed by the dropwise addition of trifluoroacetic acid (3 mL) and the reaction mixture was then allowed to stir for 2 h or until complete consumption of the starting material as indicated

by TLC. After removal of the solvent *in vacuo*, the reaction mixture was diluted with DCM (30 mL), H₂O (20 mL) and neutralised with a saturated solution of NaHCO₃ (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 12.5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **32** as an off white solid (84.6 mg, 0.163 mmol, 72%).

R_f: 0.44 (10% MeOH in DCM); IR (ATR) cm⁻¹: 3408 (m, N-H stretch), 3213 (s, O-H stretch), 3110 (m, N-H stretch), 2935 (m, C-H stretch), 2817 (m, C-H stretch), 1672 (s, C=O stretch), 1565 (m, N-H bend), 1460 (m, C-H bend), 1400 (m, C-H bend), 1342 (s, C-N stretch), 1036 (s, C-O stretch), 729 (s, N-H wag); ¹H NMR (300 MHz, CDCl₃) δ 1.33 (t, *J* = 7.6 Hz, 3H, COCH₂CH₃), 2.26 – 2.22 (m, 8H, CH₂N(CH₃)₂ and NCH₂CH₂N), 2.45 (q, *J* = 7.6 Hz, 2H, COCH₂CH₃), 2.67 (s, 3H, ArNCH₃), 2.90 (t, *J* = 5.8 Hz, 2H, NCH₂CH₂N), 3.87 (s, 3H, ArOCH₃), 5.75 (s, 2H, NCH₂O), 6.78 (s, 1H, ArH), 7.19 (d, *J* = 5.3 Hz, 1H, ArH), 7.30 – 7.25 (m, 2H, 2 × ArH), 7.56 (dd, *J* = 6.4, 2.2 Hz, 1H, ArH), 7.78 (s, 1H, ArH), 7.98 (dd, *J* = 6.4, 2.2 Hz, 1H, ArH), 8.39 (d, *J* = 5.3 Hz, 1H, ArH), 9.25 (s, 1H, ArH), 9.61 (s, 1H, ArNH), 10.13 (s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 10.6, 31.3, 44.3, 45.3, 56.0, 56.1, 57.3, 69.4, 104.7, 107.9, 109.6, 110.3, 115.2, 119.9, 121.2, 122.1, 125.9, 129.3, 127.9, 134.0, 134.5, 137.2, 144.4, 158.3, 159.3, 161.6, 173.0 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₇H₃₄N₇O₂: 488.2774; found: 488.2783.

N-(4-Fluoro-2-methoxy-5-nitrophenyl)-4-(1H-indol-3-yl)pyrimidin-2-amine (33)²⁴

A 2-neck round-bottomed flask was charged with 3-(2-chloropyrimidin-4-yl)-1*H*-indole (1.97 g, 8.59 mmol, 1 equiv.) and 4-fluoro-2-methoxy-5-nitroaniline (1.60 g, 8.59 mmol, 1 equiv.) and purged with Ar. To this was added 2-pentanol (90 mL) and the resulting suspension was placed at 105 °C. *p*-Toluenesulfonic acid monohydrate (1.96 g, 10.3 mmol, 1.2 equiv.) was added as a single portion and the reaction mixture was then allowed to stir for 4 h or until complete consumption of the starting material

as indicated by TLC. The reaction mixture was then allowed to cool and stand for 8 h which resulted in the formation of a yellow precipitate.

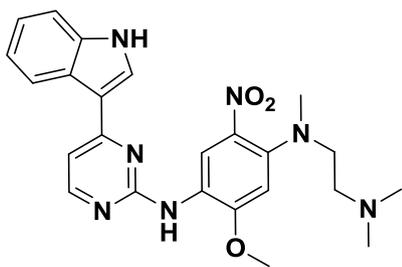
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

The precipitate was collected by filtration, washed with aliquots of acetonitrile (3×30 mL) and dried *in vacuo*. The solid was dissolved in 10% MeOH in DCM and washed with a saturated solution of NaHCO_3 (3×40 mL), a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the solid was then triturated with acetonitrile and dried under vacuum to afford compound **33** as a mustard-yellow powder (2.90 g, 7.65 mmol, 89%) with no further purification required.

^1H NMR and MS data collected for this compound compared well with the reported literature values.²⁴

Rf: 0.43 (2% MeOH in DCM); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 4.01 (s, 3H, ArOCH_3), 7.14 – 7.07 (m, 1H, ArH), 7.23 – 7.16 (m, 1H, ArH), 7.32 (d, $J = 13.4$ Hz, 1H, ArH), 7.37 (d, $J = 5.4$ Hz, 1H, ArH), 7.47 (d, $J = 8.0$ Hz, 1H, ArH), 8.22 (s, 1H, ArNH), 8.41 – 8.33 (m, 3H, $3 \times \text{ArH}$), 9.13 (d, $J = 8.5$ Hz, 1H, ArH), 11.89 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 57.4, 101.1 (d, $J = 26.3$ Hz), 108.5, 112.1, 113.2, 116., 120.7, 121.7, 122.2, 125.0, 125.9, 128.6 (d, $J = 6.7$ Hz), 129.4, 137.2, 155.3 (d, $J = 10.0$ Hz), 157.1, 159.4, 162.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{15}\text{N}_5\text{O}_3\text{F}$: 380.1159; found: 380.1148.

***N*1-[4-(1*H*-indol-3-yl)pyrimidin-2-yl]-*N*4-[2-(dimethylamino)ethyl]-2-methoxy-*N*4-methyl-5-nitrobenzene-1,4-diamine (**34**)**³⁷



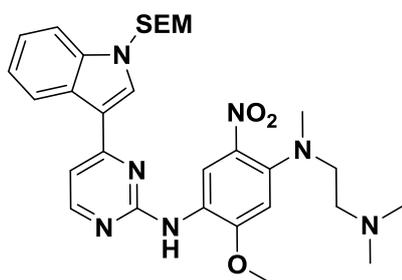
A 2-neck round-bottom flask was charged with *N*-(4-fluoro-2-methoxy-5-nitrophenyl)-4-(1*H*-indol-3-yl)pyrimidin-2-amine (3.20 g, 8.44 mmol, 1 equiv.), purged with Ar and dissolved in DMA (50 mL). To this solution was added DIPEA (2.94 mL, 16.9 mmol, 2 equiv.), followed by the slow addition of a solution of N^1,N^1,N^2 -trimethylethane-1,2-diamine (1.08 g, 10.5 mmol, 1.25 equiv.) in DMA (10 mL) at rt. The reaction was placed at 60 °C and

allowed to stir for 4 h, during which time the reaction mixture developed a deep red colour. After allowing to cool to rt, reaction mixture was then diluted with H_2O (70 mL) and EtOAc (70 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×30 mL) and the combined organic layers were then washed with successive aliquots of a saturated solution of brine (3×30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 2 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **34** as an orange powder (2.74 g, 6.83 mmol, 81%).

^1H NMR and MS data collected for this compound compared well with the reported literature values.³⁷

Rf: 0.32 (10% MeOH in DCM); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 2.17 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.50 (t, $J = 6.8$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.85 (s, 3H, ArNCH_3), 3.27 (t, $J = 6.8$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.96 (s, 3H, ArOCH_3), 6.85 (s, 1H, ArH), 7.08 (dd [app. t], $J = 7.2$ Hz, 1H, ArH), 7.18 (dd [app. t], $J = 7.2$ Hz, 1H, ArH), 7.29 (d, $J = 5.4$ Hz, 1H, ArH), 7.46 (d, $J = 8.0$ Hz, 1H, ArH), 8.08 (s, 1H, ArH), 8.38 – 8.30 (m, 3H, $3 \times \text{ArH}$), 8.67 (s, 1H, ArNH), 11.87 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 40.6, 45.4, 52.8, 56.3, 56.4, 102.1, 107.6, 112.0, 113.4, 119.2, 120.6, 121.4, 121.9, 122.1, 125.1, 129.1, 132.3, 137.1, 143.7, 154.7, 157.0, 160.1, 162.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{24}\text{H}_{28}\text{N}_7\text{O}_3$: 462.2252; found: 462.2246.

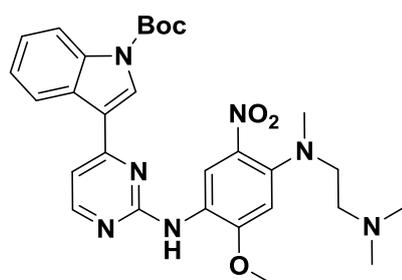
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

N1-[2-(dimethylamino)ethyl]-5-methoxy-N1-methyl-2-nitro-N4-[4-(1-[2-(trimethylsilyl)ethoxy]methyl)-1H-indol-3-yl]pyrimidin-2-yl]benzene-1,4-diamine (35)

A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in mineral oil, 26.0 mg, 0.649 mmol, 1.2 equiv.), purged with Ar and to this was added DMF (10 mL). The resulting suspension was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) and a solution of N1-[4-(1H-indol-3-yl)pyrimidin-2-yl]-N4-[2-(dimethylamino)ethyl]-2-methoxy-N4-methyl-5-nitrobenzene-1,4-diamine (250 mg, 0.542 mmol, 1 equiv.) in DMF (5 mL)

was added dropwise over 15 min. The orange solution was then allowed to warm to rt and stirred for 30 min, during which time the solution darkened. After recooling the solution to $-10\text{ }^{\circ}\text{C}$ (acetone/ice), a solution of SEM-Cl (108 mg, 0.648 mmol, 1.2 equiv.) in DMF (5 mL) was added dropwise over 10 min. The reaction mixture was then allowed to warm to rt and stirred for 4 h or until complete consumption of the starting material as indicated by TLC, changing to a light orange colour during this period. The reaction mixture was then quenched and diluted with H_2O (30 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc ($3 \times 20\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine ($3 \times 30\text{ mL}$), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **35** as an orange solid (153 mg, 0.259 mmol, 48%).

Rf: 0.45 (10% MeOH in DCM); $^1\text{H NMR}$ (300 MHz, CDCl_3) δ -0.09 (s, 9H, $\text{CH}_2\text{Si}(\text{CH}_3)_3$), 0.91 (t, $J = 8.1\text{ Hz}$, 2H, $\text{CH}_2\text{CH}_2\text{Si}$), 2.26 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.57 (t, $J = 7.0\text{ Hz}$, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.88 (s, 3H, ArNCH_3), 3.27 (t, $J = 7.0\text{ Hz}$, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.57 (t, $J = 8.1\text{ Hz}$, 2H, OCH_2CH_2), 3.93 (s, 3H, ArOCH_3), 5.59 (s, 2H, NCH_2O), 6.63 (s, 1H, ArH), 7.16 (d, $J = 5.3\text{ Hz}$, 1H, ArH), 7.32 – 7.28 (m, 2H, $2 \times \text{ArH}$), 7.53 (s, 1H, ArH), 7.57 (dd, $J = 6.4, 2.8\text{ Hz}$, 1H, ArH), 8.14 (dd, $J = 6.4, 2.8\text{ Hz}$, 1H, ArH), 8.30 (s, 1H, ArH), 8.37 (d, $J = 5.3\text{ Hz}$, 1H, ArH), 9.52 (s, 1H, ArNH) ppm; $^{13}\text{C NMR}$ (75 MHz, CDCl_3) δ -1.4, 17.7, 41.2, 45.9, 54.1, 56.1, 57.0, 66.2, 76.6, 101.7, 108.4, 111.2, 114.7, 116.3, 120.8, 121.9, 122.8, 123.1, 126.3, 132.1, 133.9, 137.5, 143.0, 152.3, 158.1, 159.3, 161.6 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{42}\text{N}_7\text{O}_4\text{Si}$: 592.3068; found: 592.3069.

tert-Butyl 3-{2-[(4-{2-(dimethylamino)ethyl}(methyl)amino)-2-methoxy-5-nitrophenyl]amino]pyrimidin-4-yl}-1H-indole-1-carboxylate (36)

A 2-neck round-bottom flask was charged with N1-[4-(1H-indol-3-yl)pyrimidin-2-yl]-N4-[2-(dimethylamino)ethyl]-2-methoxy-N4-methyl-5-nitrobenzene-1,4-diamine (2.17 g, 4.70 mmol, 1 equiv.), purged with Ar, dissolved in DMF (25 mL) and DIPEA (0.900 mL, 5.17 mmol, 1.1 equiv.) was added. After cooling to $0\text{ }^{\circ}\text{C}$, a solution of Boc_2O (1.13 g, 5.17 mmol, 1.1 equiv.) in DMF (10 mL) was added dropwise, followed by the addition of

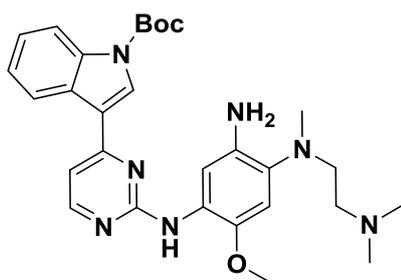
a catalytic amount of DMAP. The reaction mixture was then allowed to warm to rt and stirred for 6 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (100 mL) and EtOAc (70 mL) and the organic layer was separated.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

The aqueous layer was extracted with aliquots of EtOAc (3 × 30 mL) and the combined organic layers were then washed with successive aliquots of a saturated solution of brine (3 × 30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **36** as an orange foam (2.58 g, 4.60 mmol, 98%).

Rf: 0.45 (10% MeOH in DCM); ¹H NMR (300 MHz, CDCl₃) δ 1.71 (s, 9H, OC(CH₃)₃), 2.26 (s, 6H, CH₂N(CH₃)₂), 2.56 (t, *J* = 7.0 Hz, 2H, NCH₂CH₂N) 2.87 (s, 3H, ArNCH₃), 3.26 (t, *J* = 7.0 Hz, 2H, NCH₂CH₂N), 3.97 (s, 3H, ArOCH₃), 6.67 (s, 1H, ArH), 7.14 (d, *J* = 5.2 Hz, 1H, ArH), 7.40 – 7.35 (m, 2H, 2 × ArH), 7.57 (s, 1H, ArH), 8.25 – 8.21 (m, 1H, ArH), 8.28 (s, 1H, ArH), 8.36 – 8.31 (m, 1H, ArH), 9.19 (s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 27.6, 40.4, 45.4, 52.7, 56.3, 56.4, 84.8, 102.0, 108.6, 114.7, 118.1, 120.6, 120.7, 122.7, 123.3, 124.9, 127.3, 128.2, 131.9, 135.4, 144.3, 148.7, 155.7, 158.0, 160.4, 160.7 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₉H₃₆N₇O₅: 562.2778; found: 562.2776.

tert-Butyl 3-{2-[5-amino-4-{2-(dimethylamino)ethyl}(methyl)amino]-2-methoxyphenyl}amino]pyrimidin-4-yl}-1H-indole-1-carboxylate (37)

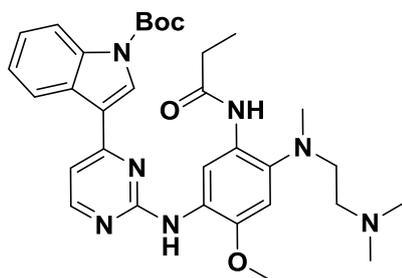


A 3-neck round-bottom flask was charged with 4-fluoro-2-methoxy-1-nitrobenzene (2.59 g, 4.61 mmol, 1 equiv.) and Pd/C (49.0 mg, 0.460 mmol, 10 mol%), placed under vacuum and MeOH (30 mL) was added. A hydrogen atmosphere was introduced *via* balloon and the suspension was then allowed to stir at rt for 12 h or until complete consumption of starting material as indicated by TLC. The reaction mixture was filtered through a plug of celite,

washed with MeOH (35 mL) and the solvent was removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **37** as a brown foam (2.11 g, 3.96 mmol, 86%).

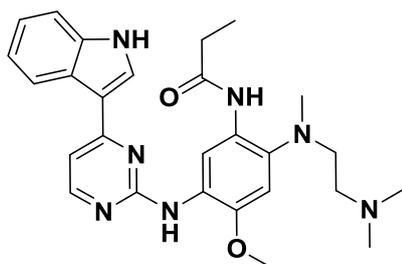
Rf: 0.38 (10% MeOH in DCM); ¹H NMR (300 MHz, DMSO-*d*₆) δ 1.67 (s, 9H, OC(CH₃)₃), 2.19 (s, 6H, CH₂N(CH₃)₂), 2.39 (t, *J* = 6.7 Hz, 2H, NCH₂CH₂N), 2.64 (s, 3H, ArNCH₃), 2.91 (t, *J* = 6.7 Hz, 2H, NCH₂CH₂N), 3.71 (s, 3H, ArOCH₃), 4.59 (br s, 2H, ArNH₂), 6.77 (s, 1H, ArH), 7.31 – 7.25 (m, 2H, 2 × ArH), 7.36 (s, 1H, ArH), 7.38 (s, 1H, ArH), 8.15 – 8.09 (m, 2H, 2 × ArH), 8.37 (d, *J* = 5.3 Hz, 1H, ArH), 8.51 (s, 1H, ArNH), 8.53 (s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 27.7, 41.5, 45.6, 53.8, 56.4, 57.3, 84.7, 105.4, 107.8, 110.6, 114.6, 118.2, 123.1, 123.4, 124.6, 124.9, 127.4, 128.0, 134.4, 135.4, 136.7, 142.9, 148.8, 157.8, 160.6, 160.8 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₉H₃₈N₇O₃: 532.3036; found: 532.3034.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

tert-Butyl 3-{2-[(4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxy-5-propionamidophenyl)amino]pyrimidin-4-yl}-1H-indole-1-carboxylate (38)

A 2-neck round-bottom flask was charged with *tert*-butyl 3-{2-[(5-amino-4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxyphenyl)amino]pyrimidin-4-yl}-1*H*-indole-1-carboxylate (2.00 g, 3.76 mmol, 1 equiv.), purged with Ar and dissolved in DCM (30 mL). The solution was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) and DIPEA (0.790 mL, 4.51 mmol, 1.2 equiv.) was added. A solution of propionyl chloride (383 mg, 4.14 mmol, 1.1 equiv.) in DCM (10 mL) was then added dropwise to the solution over 10 min. The reaction mixture was then allowed to warm to rt and stirred for 2 h. The reaction mixture was then quenched and diluted with H_2O (30 mL), a saturated solution of NaHCO_3 (30 mL) and DCM (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM ($3 \times 30\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine (50 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 12.5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **38** as a brown solid (2.10 g, 3.57 mmol, 95%).

Rf: 0.53 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3419 (m, N-H stretch), 3132 (m, N-H stretch), 2935 (m, C-H stretch), 2815 (m, C-H stretch), 1732 (s, C=O stretch), 1674 (s, C=O stretch), 1581 (m, N-H bend), 1452 (m, C-H bend), 1402 (m, C-H bend), 1357 (s, C-N stretch), 1149 (s, C-O stretch), 1054 (s, C-O stretch), 764 (s, N-H wag), 753 (s, N-H wag); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 1.05 (t, $J = 7.6\text{ Hz}$, 3H, COCH_2CH_3), 1.66 (s, 9H, $\text{OC}(\text{CH}_3)_3$), 2.24 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.30 (q, $J = 7.6\text{ Hz}$, 2H, COCH_2CH_3), 2.35 (t, $J = 5.5\text{ Hz}$, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.71 (s, 3H, ArNCH_3), 2.91 (t, $J = 5.5\text{ Hz}$, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.78 (s, 3H, ArOCH_3), 7.00 (s, 1H, *ArH*), 7.22 (dd [app t.], $J = 7.5\text{ Hz}$, 1H, *ArH*), 7.41 – 7.31 (m, 2H, $2 \times \text{ArH}$), 8.11 (d, $J = 8.3\text{ Hz}$, 1H, *ArH*), 8.33 (s, 1H, *ArH*), 8.37 (d, $J = 5.3\text{ Hz}$, 1H, *ArH*), 8.43 (d, $J = 7.5\text{ Hz}$, 1H, *ArH*), 8.52 (s, 1H, *ArH*), 8.55 (s, 1H, *ArNH*), 9.75 (s, 1H, *ArNH*), ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 9.8, 27.7, 29.5, 42.4, 45.3, 55.8, 56.9, 84.7, 105.4, 108.0, 114.6, 117.2, 118.2, 123.1, 123.2, 124.4, 124.9, 127.4, 127.6, 128.1, 135.4, 139.1, 148.1, 148.8, 157.9, 160.6, 160.8, 171.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{32}\text{H}_{42}\text{N}_7\text{O}_4$: 588.3298; found: 588.3300.

N-(5-{[4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino}-2-{[2-(dimethylamino)ethyl](methyl)amino]-4-methoxyphenyl)propionamide (27)

A 2-neck round-bottomed flask was charged with *tert*-butyl 3-{2-[(4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxy-5-propionamidophenyl)amino]pyrimidin-4-yl}-1*H*-indole-1-carboxylate (1.80 g, 3.06 mmol, 1 equiv.), purged with Ar and dissolved in DCM (18 mL). This was followed by the dropwise addition of trifluoroacetic acid (6 mL) and the reaction mixture was then allowed to stir for 4 h or until complete consumption of the starting material as indicated by TLC.

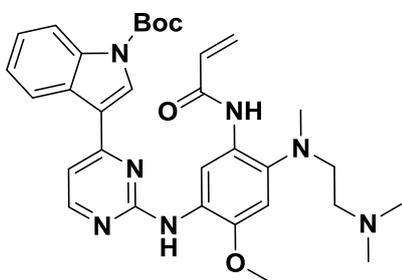
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

After removal of the solvent *in vacuo*, the reaction mixture was diluted with DCM (30 mL), H₂O (30 mL) and neutralised with a saturated solution of NaHCO₃ (60 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 30 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO₄ and filtered.

After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **27** as an off-white foam (1.43 g, 2.94 mmol, 96%).

For spectroscopic characterization data please see compound **27** on page 219/220.

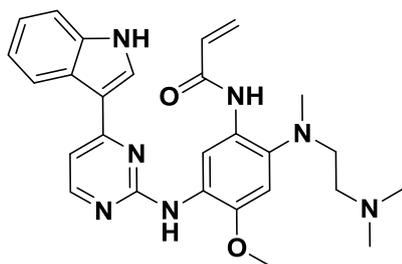
tert-Butyl 3-{2-[(5-acrylamido-4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxyphenyl)amino]pyrimidin-4-yl}-1H-indole-1-carboxylate (39)



A 2-neck round-bottom flask was charged with *tert*-butyl 3-{2-[(5-amino-4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxyphenyl)amino]pyrimidin-4-yl}-1H-indole-1-carboxylate (200 mg, 0.376 mmol, 1 equiv.), purged with Ar and dissolved in DCM (30 mL). DIPEA (80.0 μ L, 0.450 mmol, 1.2 equiv.) was added and the solution was cooled to –40 °C (acetonitrile/ CO₂). A solution of acryloyl chloride (37.0 mg, 0.409 mmol, 1.1 equiv.) in DCM (10 mL) was then added dropwise to the solution over 10 min. The reaction mixture was then allowed to warm to –20 °C and stirred at this temperature for 2 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H₂O (20 mL), a saturated solution of NaHCO₃ (20 mL) and DCM (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 20 mL) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **39** as a white foam (178 mg, 0.304 mmol, 81%).

R_f: 0.48 (10% MeOH in DCM); IR (ATR) cm⁻¹: 3423 (m, N-H stretch), 3141 (m, N-H stretch), 2949 (m, C-H stretch), 2832 (m, C-H stretch), 1738 (s, C=O stretch), 1676 (s, C=O stretch), 1581 (m, N-H bend), 1452 (m, C-H bend), 1402 (m, C-H bend), 1357 (s, C-N stretch), 1151 (s, C-O stretch), 1062 (s, C-O stretch), 970 (s, C=C bend), 764 (s, N-H wag); ¹H NMR (300 MHz, CDCl₃) δ 1.68 (s, 9H, OC(CH₃)₃), 2.26 (s, 6H, CH₂N(CH₃)₂), 2.31 (t, *J* = 5.5, 2H, NCH₂CH₂N), 2.68 (s, 3H, ArNCH₃), 2.87 (t, 2H, *J* = 5.5, NCH₂CH₂N), 3.87 (s, 3H, ArOCH₃), 5.65 (dd, *J* = 10.6, 3.2 Hz, 1H, CCH=CH₂), 6.39 – 6.35 (m, 2H, CCH=CH₂ and CCH=CH₂), 6.78 (s, 1H, ArH), 7.10 (d, *J* = 5.2 Hz, 1H, ArH), 7.36 – 7.29 (m, 2H, 2 × ArH), 7.61 (s, 1H, ArH), 8.19 (d, *J* = 7.5 Hz, 1H, ArH), 8.31 (s, 1H, ArH), 8.38 (d, *J* = 7.5 Hz, 1H, ArH), 8.50 (d, *J* = 5.2 Hz, 1H, ArH), 9.59 (s, 1H, ArNH), 10.05 (s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 28.2, 43.5, 45.4, 56.1, 56.2, 57.3, 84.5, 104.6, 109.1, 112.0, 115.2, 119.2, 122.0, 123.5, 124.8, 125.9, 126.6, 127.6, 127.8, 129.3, 132.4, 135.9, 136.2, 145.1, 149.4, 158.0, 160.2, 161.4, 163.3 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₃₂H₄₀N₇O₄: 586.3142; found: 586.3151.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

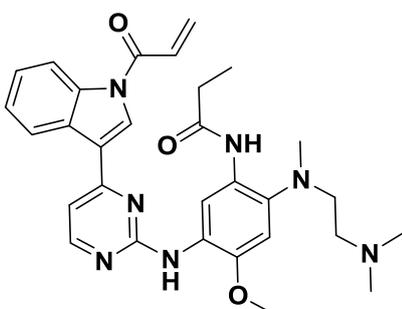
N-(5-{[4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino}-2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxyphenyl)acrylamide (HL1116 40)³⁷

A 2-neck round-bottomed flask was charged with *tert*-butyl 3-{2-[(5-acrylamido-4-{[2-(dimethylamino)ethyl](methyl)amino}-2-methoxyphenyl)amino]pyrimidin-4-yl}-1*H*-indole-1-carboxylate (140 mg, 0.239 mmol, 1 equiv.), purged with Ar and dissolved in DCM (12 mL). This was followed by the dropwise addition of trifluoroacetic acid (4 mL) and the reaction mixture was then allowed to stir for 6 h or until complete

consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, the reaction mixture was diluted with DCM (30 mL), H₂O (30 mL) and neutralised with a saturated solution of NaHCO₃ (60 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 30 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **40** as a white foam (81.0 mg, 0.167 mmol, 70%).

¹H NMR and MS data collected for this compound compared well with the reported literature values.³⁷

Rf: 0.23 (10% MeOH in DCM); ¹H NMR (300 MHz, DMSO-*d*₆) δ 2.25 (s, 6H, CH₂N(CH₃)₂), 2.37 (br s, 2H, NCH₂CH₂N), 2.71 (s, 3H, ArNCH₃), 2.92 (br s, 2H, NCH₂CH₂N), 3.84 (s, 3H, ArOCH₃), 5.75 (d, *J* = 8.1 Hz, 1H, CCH=CH₂), 6.28 (d, *J* = 16.2 Hz, 1H, CCH=CH₂), 6.46 (dd, *J* = 16.2, 10.2 Hz, 1H, CCH=CH₂), 7.02 (s, 1H, ArH), 7.11 – 7.04 (m, 1H, ArH), 7.20 – 7.12 (m, 1H, ArH), 7.26 (d, *J* = 5.0 Hz, 1H, ArH), 7.45 (d, *J* = 7.8 Hz, 1H, ArH), 7.96 (s, 1H, ArH), 8.29 (s, 1H, ArH), 8.30 (s, 1H, ArH), 8.50 (s, 1H, ArH), 8.97 (s, 1H, ArH), 10.14 (s, 1H, ArNH), 11.81 (s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 42.6, 45.0, 55.9, 56.7, 105.2, 107.2, 111.9, 113.5, 115.1, 120.5, 121.8, 121.9, 125.1, 125.2, 125.2, 126.2, 127.4, 129.5, 132.3, 137.2, 138.1, 146.8, 157.2, 160.2, 162.4, 162.5 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₇H₃₂N₇O₂: 486.2617; found: 486.2622.

N-(5-{[4-(1-Acryloyl-1*H*-indol-3-yl)pyrimidin-2-yl]amino}-2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxyphenyl)propionamide (41)

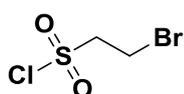
A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in mineral oil, 5.20 mg, 0.134 mmol, 1.3 equiv.), purged with Ar and to this was added DMF (4 mL). The resulting suspension was cooled to –10 °C (acetone/ice) and a solution of *N*-(5-{[4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino}-2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxyphenyl)propionamide (50.0 mg, 0.103 mmol, 1 equiv.) in DMF (1 mL) was added dropwise over 5 min. The solution was then allowed to

warm to rt and stirred for 30 min, during which time the colour darkened.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

After recooling the solution to $-40\text{ }^{\circ}\text{C}$ (acetonitrile/ CO_2), a solution of acryloyl chloride (11.1 mg, 0.124 mmol, 1.2 equiv.) in DMF (1 mL) was added dropwise over 5 min. The reaction mixture was then allowed to warm to $-20\text{ }^{\circ}\text{C}$ and stirred at this temperature for 2 h or until complete consumption of the starting material as indicated by TLC, the colour changing from dark to light brown during this period. The reaction mixture was then quenched and diluted with H_2O (30 mL), a saturated solution of NaHCO_3 (20 mL) and EtOAc (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc ($3 \times 20\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine ($3 \times 30\text{ mL}$), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **41** as yellow solid (40.0 mg, 73.9 μmol , 72%).

Rf: 0.37 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3453 (m, N-H stretch), 3149 (m, N-H stretch), 2937 (m, C-H stretch), 2852 (m, C-H stretch), 1695 (s, C=O stretch), 1666 (s, C=O stretch), 1579 (m, N-H bend), 1452 (m, C-H bend), 1402 (m, C-H bend), 1349 (s, C-N stretch), 1155 (s, C-O stretch), 1021 (s, C-O stretch), 976 (s, C=C bend), 768 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 1.26 (t, $J = 7.6\text{ Hz}$, 3H, COCH_2CH_3), 2.36 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.53 – 2.40 (m, 4H, $\text{NCH}_2\text{CH}_2\text{N}$ and COCH_2CH_3), 2.70 (s, 3H, ArNCH_3), 2.98 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.89 (s, 3H, ArOCH_3), 6.06 (dd, $J = 16.2, 10.2\text{ Hz}$, 1H, $\text{CCH}=\text{CH}_2$), 6.81 – 6.69 (m, 2H, $\text{CCH}=\text{CH}_2$ and $\text{CCH}=\text{CH}_2$), 7.24 (d, $J = 5.1\text{ Hz}$, 1H, ArH), 7.26 (s, 1H, ArH), 7.47 – 7.35 (m, 2H, $2 \times \text{ArH}$), 7.74 (s, 1H, ArH), 7.92 – 7.81 (m, 1H, ArH), 8.03 (d, $J = 7.2\text{ Hz}$, 1H, ArH), 8.50 (d, $J = 5.1\text{ Hz}$, 1H, ArH), 8.68 (d, $J = 7.2\text{ Hz}$, 1H, ArH), 8.91 (s, 1H, ArH), 9.68 (s, 1H, ArNH), 9.75 (s, 1H, ArNH), ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 10.6, 31.1, 43.3, 44.1, 45.3, 56.3, 57.2, 104.5, 110.2, 110.4, 115.1, 117.6, 120.2, 121.3, 122.1, 124.6, 125.5, 127.0, 127.8, 128.2, 129.1, 129.6, 132.2, 137.3, 144.3, 158.6, 160.2, 160.8, 165.0 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{36}\text{N}_7\text{O}_3$: 542.2880; found: 542.2881.

2-Bromoethane-1-sulfonyl chloride (46)²⁸

A flame-dried, 2-neck round-bottomed flask was charged with sodium 2-bromoethane-1-sulfonate (6.33 g, 30.0 mmol, 1 equiv.) and purged with Ar. To this was added PCl_5 (6.24 g, 30.0 mmol, 1 equiv.) portion wise over 5 min with vigorous stirring which resulted in an exotherm and the reaction mixture becoming a yellow slurry. The reaction mixture was then allowed to stir at $130\text{ }^{\circ}\text{C}$ for 30 minutes, followed by an additional 30 minutes at $70\text{ }^{\circ}\text{C}$. After allowing to cool to rt, the reaction mixture was slowly poured into vigorously stirring H_2O (80 mL) and allowed to stir for 10 minutes. The reaction mixture was diluted with DCM (40 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM ($3 \times 40\text{ mL}$) and the combined organic layers were then washed with H_2O (50 mL), an ice-cold solution of 5% NaHCO_3 (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved by kugelrohr distillation. Fractions with boiling point between $135\text{ }^{\circ}\text{C}$ and $145\text{ }^{\circ}\text{C}$ at $\approx 20\text{ mbar}$ were collected to afford compound **46** as a clear liquid (5.04 g, 24.3 mmol, 81%).

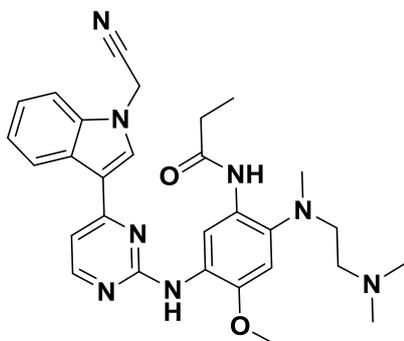
^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.²⁸

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

^1H NMR (300 MHz, CDCl_3) δ 3.81 – 3.74 (m, 2H, $\text{BrCH}_2\text{CH}_2\text{SO}_2\text{Cl}$), 4.15 – 4.07 (m, 2H, $\text{BrCH}_2\text{CH}_2\text{SO}_2\text{Cl}$) ppm;

^{13}C NMR (75 MHz, CDCl_3) δ 20.11, 65.88 ppm.

***N*-[5-({4-[1-(Cyanomethyl)-1*H*-indol-3-yl]pyrimidin-2-yl}amino)-2-{[2-(dimethylamino)ethyl](methylamino)-4-methoxyphenyl]propionamide (53)**

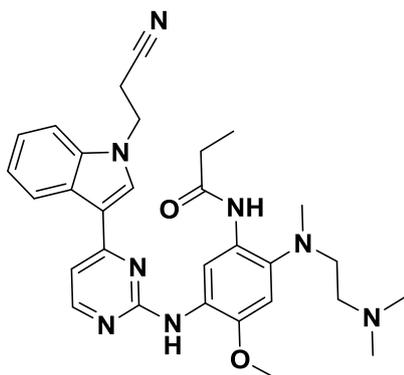


A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in mineral oil, 10.7 mg, 0.267 mmol, 1.3 equiv.), purged with Ar and to this was added DMF (4 mL). The resulting suspension was cooled to $-10\text{ }^\circ\text{C}$ (acetone/ice) and a solution of *N*-[5-({4-(1*H*-indol-3-yl)pyrimidin-2-yl}amino)-2-{[2-(dimethylamino)ethyl](methylamino)-4-methoxyphenyl]propionamide (100 mg, 0.205 mmol, 1 equiv.) in DMF (1 mL) was added dropwise over 5 min. The solution was then allowed to warm to rt and stirred for 30 min, during which time the colour darkened.

After recooling the solution to $-10\text{ }^\circ\text{C}$ (acetone/ice), a solution of iodoacetonitrile (41.1 mg, 0.246 mmol, 1.2 equiv.) in DMF (1 mL) was added dropwise over 5 min. The reaction mixture was then allowed to warm to rt and stirred at this temperature for 2 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (30 mL), a saturated solution of NaHCO_3 (20 mL) and EtOAc (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc ($3 \times 20\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine ($3 \times 30\text{ mL}$), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **53** as yellow solid (73.0 mg, 0.139 mmol, 68%).

R_f: 0.40 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3402 (m, N-H stretch), 3118 (m, N-H stretch), 2912 (m, C-H stretch), 2819 (m, C-H stretch), 2772 (m, C-H stretch), 2236 (w, $\text{C}\equiv\text{N}$ stretch), 1662 (s, $\text{C}=\text{O}$ stretch), 1579 (m, N-H bend), 1464 (m, C-H bend), 1396 (m, C-H bend), 1342 (s, C-N stretch), 1013 (s, C-O stretch), 731 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 1.31 (t, $J = 7.6\text{ Hz}$, 3H, COCH_2CH_3), 2.28 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.31 (t, $J = 5.8\text{ Hz}$, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.48 (q, $J = 7.6\text{ Hz}$, 2H, COCH_2CH_3), 2.68 (s, 3H, ArNCH_3), 2.92 (t, $J = 5.8\text{ Hz}$, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.88 (s, 3H, ArOCH_3), 5.44 (s, 2H, $\text{NCH}_2\text{C}\equiv\text{N}$), 6.78 (s, 1H, ArH), 7.18 (d, $J = 5.3\text{ Hz}$, 1H, ArH), 7.37 – 7.30 (m, 2H, $2 \times \text{ArH}$), 7.48 (dd, $J = 6.6, 2.2\text{ Hz}$, 1H, ArH), 7.78 (s, 1H, ArH), 8.03 (dd, $J = 6.6, 2.2\text{ Hz}$, 1H, ArH), 8.39 (d, $J = 5.2\text{ Hz}$, 1H, ArH), 9.37 (s, 1H, ArH), 9.58 (s, 1H, ArNH), 9.87 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 10.3, 30.8, 34.4, 44.3, 45.4, 55.8, 56.2, 57.3, 104.7, 108.0, 109.8, 110.0, 114.6, 115.9, 120.7, 122.1, 123.0, 126.3, 127.6, 129.5, 134.5, 134.7, 137.1, 144.2, 158.6, 159.4, 161.1, 171.9 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{29}\text{H}_{35}\text{N}_8\text{O}_2$: 527.2883; found: 527.2889.

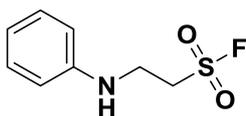
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

N-[5-({4-[1-(2-cyanoethyl)-1H-indol-3-yl]pyrimidin-2-yl}amino)-2-{{2-(dimethylamino)ethyl}(methyl)amino}-4-methoxyphenyl]propionamide (56)

A 2-neck round-bottomed flask was charged with *N*-(5-{{4-[1-(2-cyanoethyl)-1H-indol-3-yl]pyrimidin-2-yl}amino)-2-{{2-(dimethylamino)ethyl}(methyl)amino}-4-methoxyphenyl)propionamide (40.0 mg, 82.0 μ mol, 1 equiv.) and purged with Ar. To this was added DMF (4 mL) and the resulting suspension cooled to 0 °C, followed by the dropwise addition of a solution of acrylonitrile (17.4 mg, 0.328 mmol, 4 equiv.) in DMF (1 mL) and a catalytic amount of Triton B. The reaction mixture was then allowed to warm to rt and stirred for 6 h or until complete consumption of the starting material as indicated by

TLC. The reaction mixture was then diluted with H₂O (30 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 \times 20 mL) and the combined organic layers were then washed with a saturated solution of brine (3 \times 30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **56** as an off white solid (36.0 mg, 66.6 μ mol, 81%).

Rf: 0.42 (10% MeOH in DCM); IR (ATR) cm⁻¹: 3417 (m, N-H stretch), 3116 (m, N-H stretch), 2937 (m, C-H stretch), 2851 (m, C-H stretch), 2203 (w, C \equiv N stretch), 1666 (s, C=O stretch), 1577 (m, N-H bend), 1464 (m, C-H bend), 1394 (m, C-H bend), 1347 (s, C-N stretch), 1023 (s, C-O stretch), 739 (s, N-H wag); ¹H NMR (300 MHz, CDCl₃) δ 1.34 (t, *J* = 7.6 Hz, 3H, COCH₂CH₃), 2.31 (s, 6H, CH₂N(CH₃)₂), 2.36 (br s, 2H, NCH₂CH₂N), 2.51 (q, *J* = 7.6 Hz, 2H, COCH₂CH₃), 2.69 (s, 3H, ArNCH₃), 2.96 (t, *J* = 5.6 Hz, 2H, NCH₂CH₂N), 3.10 (t, *J* = 7.6 Hz, 2H, NCH₂CH₂C \equiv N), 3.88 (s, 3H, ArOCH₃), 4.71 (t, *J* = 7.6 Hz, 2H, NCH₂CH₂C \equiv N), 6.78 (s, 1H, ArH), 7.22 (d, *J* = 5.2 Hz, 1H, ArH), 7.35 – 7.27 (m, 2H, 2 \times ArH), 7.43 (d, *J* = 6.8 Hz, 1H, ArH), 7.77 (s, 1H, ArH), 8.04 (d, *J* = 6.8 Hz, 1H, ArH), 8.40 (d, *J* = 5.2 Hz, 1H, ArH), 9.31 (s, 1H, ArH), 9.67 (s, 1H, ArNH), 9.79 (s, 1H, ArNH) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 10.7, 19.0, 31.2, 42.1, 44.3, 45.4, 55.7, 56.3, 57.3, 104.7, 108.1, 109.4, 109.7, 115.5, 117.4, 120.6, 121.6, 122.5, 126.1, 127.7, 129.7, 133.7, 134.1, 137.1, 144.0, 158.4, 159.6, 161.5, 172.0 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₃₀H₃₇N₈O₂: 541.3039; found: 541.3034.

2-(Phenylamino)ethane-1-sulfonyl fluoride (62)³²

A 2-neck round-bottomed flask was charged with aniline (250 mg, 2.68 mmol, 1 equiv.) and purged with Ar. To this was added DMF (10 mL) and the resulting solution was then cooled to 0 °C, followed by the dropwise addition of a solution of ethenesulfonyl fluoride

(310 mg, 2.81 mmol, 1.05 equiv.) in DMF (5 mL). The reaction mixture was then allowed to warm to rt and stirred for 3 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H₂O (30 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 \times 20 mL) and the combined organic layers were then washed with a saturated solution of brine (3 \times 30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 20% EtOAc in PE.

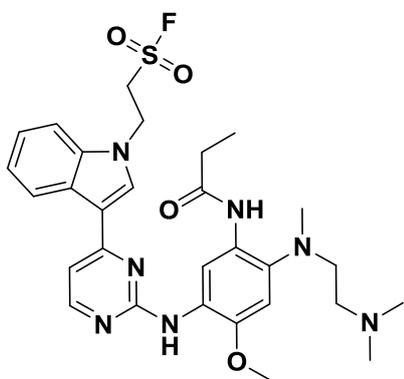
Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

Pure fractions were evaporated to dryness to afford compound **62** as a brown oil (535 mg, 2.63 mmol, 98%).

IR and ^1H NMR data collected for this compound compared well with the reported literature values.³²

Rf: 0.58 (30% PE in EtOAc); ^1H NMR (300 MHz, CDCl_3) δ 3.63 (m [app. dd], $J = 11.2, 5.5$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{S}$), 3.82 (t, $J = 6.2$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{S}$), 4.10 (br s, 1H, ArNHCH_2), 6.71 – 6.63 (m, 2H, $2 \times \text{ArH}$), 6.87 (dd [app. t], $J = 7.2$ Hz, 1H, ArH), 7.34 – 7.24 (m, 2H, $2 \times \text{ArH}$) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 38.1, 49.8 (d, $J = 13.7$ Hz), 113.2, 119.0, 129.7, 145.7 ppm; ^{19}F NMR (282 MHz, CDCl_3) δ 56.70 (t, $J = 4.6$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_8\text{H}_{11}\text{NO}_2\text{SF}$: 204.0495; found: 204.0489.

2-(3-{2-[(4-{2-(Dimethylamino)ethyl}(methyl)amino)-2-methoxy-5-propionamidophenyl]amino]pyrimidin-4-yl}-1H-indol-1-yl)ethane-1-sulfonyl fluoride (63)



A 2-neck round-bottomed flask was charged with *N*-(5-{[4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino}-2-{[2-(dimethylamino)ethyl](methyl)amino]-4-methoxyphenyl}propionamide (80.0 mg, 0.164 mmol, 1 equiv.) and potassium carbonate (34.0 mg, 0.246 mmol, 1.5 equiv.) and purged with Ar. To this was added DMF (8 mL) and the resulting suspension cooled to 0 °C, followed by the dropwise addition of a solution of ethenesulfonyl fluoride (23.5 mg, 0.213 mmol, 1.3 equiv.) in DMF (2 mL). The reaction mixture was then allowed to warm to rt and stirred for 6 h or until complete consumption

of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (30 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (3×20 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound (HL1238 **63**) as an off white solid (62.0 mg, 0.103 mmol, 63%).

Rf: 0.47 (15% MeOH in DCM); IR (ATR) cm^{-1} : 3419 (m, N-H stretch), 3116 (m, N-H stretch), 2961 (m, C-H stretch), 2860 (m, C-H stretch), 1660 (s, C=O stretch), 1577 (m, N-H bend), 1466 (m, C-H bend), 1396 (m, C-H bend), 1351 (s, C-N stretch), 1161 (s, S=O stretch), 1025 (s, C-O stretch), 743 (s, N-H wag); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 1.13 (t, $J = 7.6$ Hz, 3H, COCH_2CH_3), 2.35 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.43 – 2.36 (m, 4H, $\text{NCH}_2\text{CH}_2\text{N}$ and COCH_2CH_3), 2.68 (s, 3H, ArNCH_3), 2.99 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.85 (s, 3H, ArOCH_3), 4.69 – 4.63 (m, 2H, $\text{NCH}_2\text{CH}_2\text{S}$), 4.89 (t, $J = 6.4$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{S}$), 7.00 (s, 1H, ArH), 7.19 (dd [app. t], $J = 7.5$ Hz, 1H, ArH), 7.23 (d, $J = 5.2$ Hz, 1H, ArH), 7.28 (dd [app. t], $J = 7.5$ Hz, 1H, ArH), 7.65 (d, $J = 8.0$ Hz, 1H, ArH), 7.94 (s, 1H, ArH), 8.28 (d, $J = 8.0$ Hz, 1H, ArH), 8.36 (d, $J = 5.2$ Hz, 1H, ArH), 8.70 (s, 1H, ArH), 8.94 (s, 1H, ArNH), 9.78 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO}-d_6$) δ 10.0, 29.7, 42.9, 44.7, 49.5, 49.7, 56.0, 56.3, 105.2, 107.3, 110.7, 113.7, 114.1, 121.2, 121.6, 122.4, 125.2, 125.4, 127.7, 132.7, 136.8, 137.3, 145.9, 156.0, 157.9, 161.3, 171.5 ppm; ^{19}F NMR (282 MHz, $\text{DMSO}-d_6$) δ 57.69 (t, $J = 6.1$ Hz); HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{29}\text{H}_{37}\text{N}_7\text{O}_4\text{SF}$: 598.2612; found: 598.2620.

6.9 References

1. Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame and J. Taunton, *Journal of the American Chemical Society*, 2017, **139**, 680-685.
2. A. M. Aronov and M. A. Murcko, *Journal of Medicinal Chemistry*, 2004, **47**, 5616-5619.
3. H. Shi, C.-J. Zhang, G. Y. J. Chen and S. Q. Yao, *Journal of the American Chemical Society*, 2012, **134**, 3001-3014.
4. S.-J. Zhu, P. Zhao, J. Yang, R. Ma, X.-E. Yan, S.-Y. Yang, J.-W. Yang and C.-H. Yun, *Oncotarget*, 2018, **9**, 13652-13665.
5. A. Narayanan and L. H. Jones, *Chemical Science*, 2015, **6**, 2650-2659.
6. J. Dong, L. Krasnova, M. G. Finn and K. B. Sharpless, *Angewandte Chemie International Edition*, 2014, **53**, 9430-9448.
7. J. Pettinger, K. Jones and M. D. Cheeseman, *Angewandte Chemie International Edition*, 2017, **56**, 15200-15209.
8. P. K. Pal, R. J. Reischer, W. J. Wechter and R. F. Colman, *Journal of Biological Chemistry*, 1978, **253**, 6644-6646.
9. N. P. Grimster, S. Connelly, A. Baranczak, J. Dong, L. B. Krasnova, K. B. Sharpless, E. T. Powers, I. A. Wilson and J. W. Kelly, *Journal of the American Chemical Society*, 2013, **135**, 5656-5668.
10. S. Choi, S. Connelly, N. Reixach, I. A. Wilson and J. W. Kelly, *Nature Chemical Biology*, 2009, **6**, 133-139.
11. T.-L. Ho, *Chemical Reviews*, 1975, **75**, 1-20.
12. U. P. Dahal, A. M. Gilbert, R. S. Obach, M. E. Flanagan, J. M. Chen, C. Garcia-Irizarry, J. T. Starr, B. Schuff, D. P. Uccello and J. A. Young, *MedChemComm*, 2016, **7**, 864-872.
13. E. Anscombe, E. Meschini, R. Mora-Vidal, Mathew P. Martin, D. Staunton, M. Geitmann, U. H. Danielson, Will A. Stanley, Lan Z. Wang, T. Reuillon, Bernard T. Golding, C. Cano, David R. Newell, Martin E. M. Noble, Stephen R. Wedge, Jane A. Endicott and Roger J. Griffin, *Chemistry & Biology*, 2015, **22**, 1159-1164.
14. F. F. Fleming, L. Yao, P. C. Ravikumar, L. Funk and B. C. Shook, *Journal of Medicinal Chemistry*, 2010, **53**, 7902-7917.
15. M. J. Boyd, S. N. Crane, J. Robichaud, J. Scheigetz, W. C. Black, N. Chauret, Q. Wang, F. Massé and R. M. Oballa, *Bioorganic & Medicinal Chemistry Letters*, 2009, **19**, 675-679.
16. R. M. Oballa, J.-F. Truchon, C. I. Bayly, N. Chauret, S. Day, S. Crane and C. Berthelette, *Bioorganic & Medicinal Chemistry Letters*, 2007, **17**, 998-1002.
17. C. Laurence, K. A. Brameld, J. Graton, J.-Y. Le Questel and E. Renault, *Journal of Medicinal Chemistry*, 2009, **52**, 4073-4086.
18. Z. Zhao and P. E. Bourne, *Drug Discovery Today*, 2018, **23**, 727-735.
19. A. Erian, S. Sherif and H. Gaber, *Molecules*, 2003, **8**, 793-865.
20. K. Uchibori, N. Inase, M. Araki, M. Kamada, S. Sato, Y. Okuno, N. Fujita and R. Katayama, *Nature Communications*, 2017, **8**, 14768-14784.
21. J. Jang, J. B. Son, C. To, M. Bahcall, S. Y. Kim, S. Y. Kang, M. Mushajiang, Y. Lee, P. A. Jänne, H. G. Choi and N. S. Gray, *European Journal of Medicinal Chemistry*, 2017, **136**, 497-510.
22. Y. Yosaatmadja, S. Silva, J. M. Dickson, A. V. Patterson, J. B. Smaill, J. U. Flanagan, M. J. McKeage and C. J. Squire, *Journal of Structural Biology*, 2015, **192**, 539-544.
23. H. Zhang, W. Wu, C. Feng, Z. Liu, E. Bai, X. Wang, M. Lei, H. Cheng, H. Feng, J. Shi, J. Wang, Z. Zhang, T. Jin, S. Chen, S. Hu and Y. Zhu, *European Journal of Medicinal Chemistry*, 2017, **135**, 12-23.
24. R. A. Ward, M. J. Anderton, S. Ashton, P. A. Bethel, M. Box, S. Butterworth, N. Colclough, C. G. Chorley, C. Chuaqui, D. A. E. Cross, L. A. Dakin, J. É. Debreczeni, C. Eberlein, M. R. V. Finlay, G. B. Hill, M. Grist, T. C. M. Klinowska, C. Lane, S. Martin, J. P. Orme, P. Smith, F. Wang and M. J. Waring, *Journal of Medicinal Chemistry*, 2013, **56**, 7025-7048.
25. G. Zhu, X. Wang, F. Wang, Y. Mao and H. Wang, *Journal of Heterocyclic Chemistry*, 2017, **54**, 2898-2901.
26. J. F. Bunnett, *Quarterly Reviews, Chemical Society*, 1958, **12**, 1-16.
27. L. Lucescu, A. Ghinet, D. Belei, B. Rigo, J. Dubois and E. Bîcu, *Bioorganic & Medicinal Chemistry Letters*, 2015, **25**, 3975-3979.
28. F. Gao, X. Yan, O. Zahr, A. Larsen, K. Vong and K. Auclair, *Bioorganic & Medicinal Chemistry Letters*, 2008, **18**, 5518-5522.
29. Y. Tu, X. Zeng, H. Wang and J. Zhao, *Organic Letters*, 2018, **20**, 280-283.
30. S. Zeeli, T. Weill, E. Finkin-Groner, C. Bejar, M. Melamed, S. Furman, M. Zhenin, A. Nudelman and M. Weinstock, *Journal of Medicinal Chemistry*, 2018, **61**, 4004-4019.

Chapter 6 – Osimertinib-derived Inhibitors Targeting the EGFR Catalytic Lysine Residue Lys745

31. O. P. J. van Linden, M. Wijnmans, L. Roumen, L. Rotteveel, R. Leurs and I. J. P. de Esch, *Journal of Organic Chemistry*, 2012, **77**, 7355-7363.
32. J. J. Krutak, R. D. Burpitt, W. H. Moore and J. A. Hyatt, *Journal of Organic Chemistry*, 1979, **44**, 3847-3858.
33. Q. Chen, P. Mayer and H. Mayr, *Angewandte Chemie International Edition*, 2016, **55**, 12664-12667.
34. P. K. Chinthakindi, K. B. Govender, A. S. Kumar, H. G. Kruger, T. Govender, T. Naicker and P. I. Arvidsson, *Organic Letters*, 2017, **19**, 480-483.
35. M. T. Morgan, A. M. McCallum and C. J. Fahrni, *Chemical Science*, 2016, **7**, 1468-1473.
36. M. Warmuth, S. Kim, X.-j. Gu, G. Xia and F. Adrián, *Current Opinion in Oncology*, 2007, **19**, 55-60.
37. M. R. V. Finlay, M. Anderton, S. Ashton, P. Ballard, P. A. Bethel, M. R. Box, R. H. Bradbury, S. J. Brown, S. Butterworth, A. Campbell, C. Chorley, N. Colclough, D. A. E. Cross, G. S. Currie, M. Grist, L. Hassall, G. B. Hill, D. James, M. James, P. Kemmitt, T. Klinowska, G. Lamont, S. G. Lamont, N. Martin, H. L. McFarland, M. J. Mellor, J. P. Orme, D. Perkins, P. Perkins, G. Richmond, P. Smith, R. A. Ward, M. J. Waring, D. Whittaker, S. Wells and G. L. Wrigley, *Journal of Medicinal Chemistry*, 2014, **57**, 8249-8267.
38. R. Dasari, A. De Carvalho, D. C. Medellin, K. N. Middleton, F. Hague, M. N. M. Volmar, L. V. Frolova, M. F. Rossato, J. J. De La Chapa, N. F. Dybdal-Hargreaves, A. Pillai, V. Mathieu, S. Rogelj, C. B. Gonzales, J. B. Calixto, A. Evidente, M. Gautier, G. Munirathinam, R. Glass, P. Burth, S. C. Pelly, W. A. L. van Otterlo, R. Kiss and A. Kornienko, *ChemMedChem*, 2015, **10**, 2014-2026.
39. N. Shindo, H. Fuchida, M. Sato, K. Watari, T. Shibata, K. Kuwata, C. Miura, K. Okamoto, Y. Hatsuyama, K. Tokunaga, S. Sakamoto, S. Morimoto, Y. Abe, M. Shiroishi, J. M. M. Caaveiro, T. Ueda, T. Tamura, N. Matsunaga, T. Nakao, S. Koyanagi, S. Ohdo, Y. Yamaguchi, I. Hamachi, M. Ono and A. Ojida, *Nature Chemical Biology*, 2019, **15**, 250-258.
40. H. Liu, Y. Lv, Y. Li, J. Cai, J. Chen, Y. Qin and M. Ji, *Journal of Chemical Research*, 2015, **39**, 318-320.

Chapter 7

Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Abstract

In this chapter, our research focus shifted from targeting the EGFR catalytic lysine residue to the development of inhibitors containing functional groups able to either reversibly interact or covalently modify the mutated Ser797 residue. As this mutant serine residue contains a harder, more inert hydroxyl nucleophile, we implemented a strategy of including electrophiles known to interact and irreversibly inhibit serine-containing enzymes. These included the use of Michael acceptors, sulfonyl fluorides, fluorophosphonates and phosphonate esters, boronic esters and acids, trifluoromethyl ketones, α -, β -haloketones, imidazoles and nitriles. Additionally, we envisaged the synthesis of dual warhead containing compounds, based on the premise of the catalytic triad in serine proteases. Building on the synthetic achievements of the previous chapter, these electrophiles were incorporated into the previously implemented osimertinib-derived scaffold on the aniline conventionally derivatised with the acrylamide Michael acceptor. The synthesis of this heterocyclic scaffold, our chosen electrophiles and the incorporation thereof are comprehensively discussed. Synthetic highlights included the unexpected development of original synthetic methodology, synthesis of novel electrophiles and effective use of multi-nuclear NMR spectroscopy as structural confirmation technique. This resulted in the synthesis of 14 final compounds which were assessed in both a biochemical and cellular setting. Amongst overall positive and encouraging results, this evaluation provided compelling trends and contrasts in activity which were successfully rationalised. Lastly, we discussed the future development of these inhibitors in overcoming the EGFR-C797S resistance mutation.

7.1 Introduction

In the previous two chapters, our research efforts extended to the covalent modification and targeted inhibition of the EGFR catalytic lysine residue (Lys745). In the body of work presented in this chapter, our focus shifts towards targeting the hydroxyl nucleophile of the mutated serine residue (Ser797) which is present in the clinically relevant EGFR-C797S mutant. As discussed in **Section 2.4.3**, this mutation is proposed to be the main mechanism of resistance towards 3rd generation inhibitors such as osimertinib.¹ In order to overcome resistance mediated by the T790M mutation, 3rd generation inhibitors incorporated an acrylamide electrophile in an appropriate position to covalently modify the Cys797 residue, as displayed in **Figure 7.1a**. The premise of this irreversible inhibition lies in the nucleophilicity of the soft thiol functionality and the reactivity-tuned acrylamide moiety, allowing for the desired Michael addition reaction to take place.² With the advent of the EGFR-C797S mutant, point mutation of the nucleophilic Cys797 covalent anchor to that of a serine residue resulted in the termination and inability of these compounds to undergo covalent bond formation under physiological conditions, as illustrated in **Figure 7.1b**. This inability stems primarily from the harder and less nucleophilic nature of the serine residue and the now inappropriate electrophile installed on these inhibitors.

As with the preceding chapters, our main aspiration was the reinstatement of irreversible inhibition through covalent bond formation, coupled with exploration of reversible interactions with Ser797. To do this, we envisioned alteration of the acrylamide group of a 3rd generation inhibitor, such as osimertinib, to a more suitably reactive electrophile or functional group which would be able to covalently modify the less nucleophilic Ser797.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

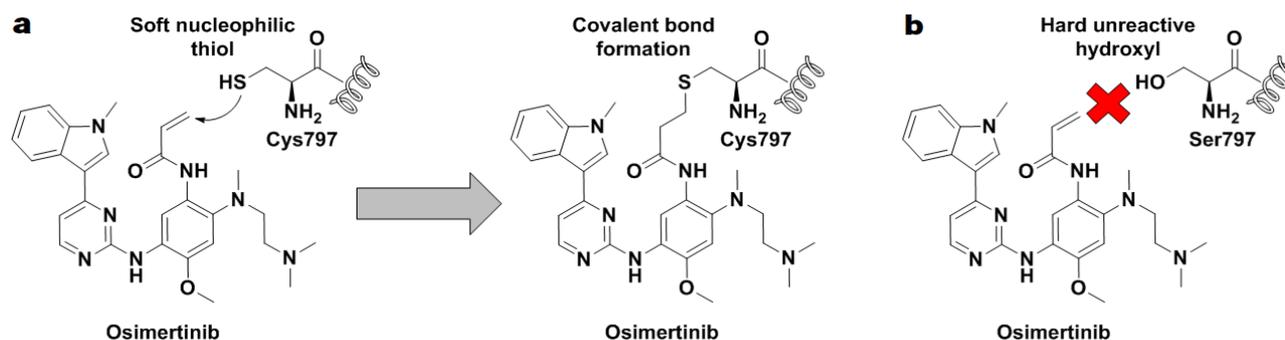


Figure 7.1: a) Irreversible inhibition through a Michael addition reaction between the acrylamide electrophile and Cys797 residue and b) No covalent bond formation as harder, less nucleophilic serine residue cannot undergo Michael addition.

We believed several inherent factors could enhance the probability of success with the implementation of this strategy. Firstly, use of 3rd generation inhibitor scaffolds, which were specifically developed with a superior and potent reversible binding affinity, would negate the requirement for preliminary driving group optimization. Furthermore, the ideal position and placement for electrophiles to target the mutated Ser797 residue would already be established, allowing us to bypass any molecular modelling and docking simulations. Thus, all that remained for us was to peruse the literature and identify the suitably reactive electrophiles and functionalities that can undergo covalent modification or favourably interact with serine residues.

7.2 Targeting Serine Residues

7.2.1 The Catalytic Triad

A catalytic triad is a combination of three concerted amino acids present in the active site of certain enzymes.³ Catalytic triads are prevalent in hydrolase and transferase enzymes, such as proteases, and typically consist of an Acid-Base-Nucleophile motif, shown in **Figure 7.2a**. This catalytic machinery aids in activation of the nucleophilic potential of specific residues, enabling their participation in covalent catalysis. As illustrated in **Figure 7.2b**, the composition of a specific enzyme's triad can differ, built from an array of amino acids in varying combinations. Keeping in line with our research topic, we will only discuss the catalytic triad where serine acts as nucleophile.

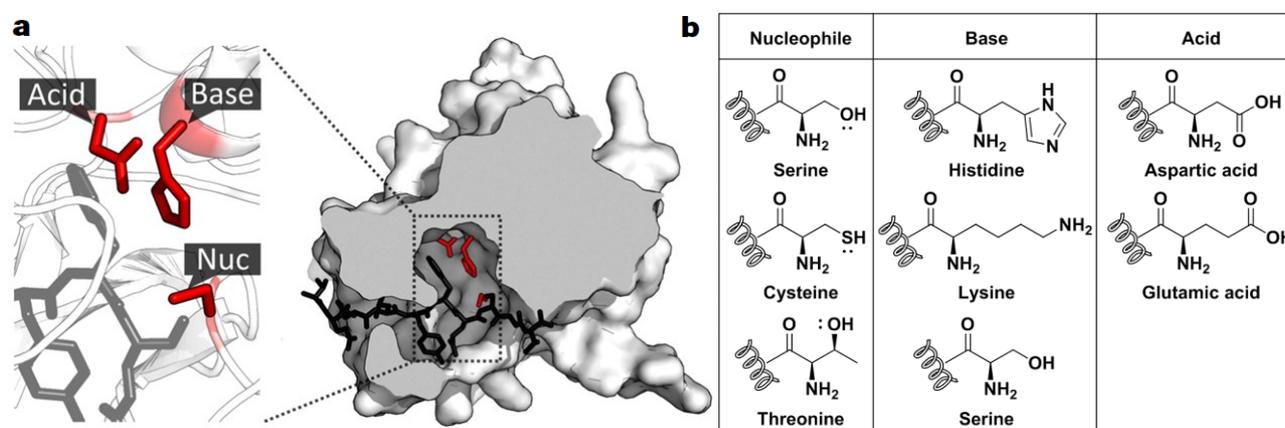


Figure 7.2: a) An example of a catalytic triad present in the active site of the Tobacco Etch Virus (TEV) protease (PDB: 1LVM) and b) Amino acid combinations constituting the catalytic triad of different enzymes. Figure reproduced from a) Phan et al.⁴

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

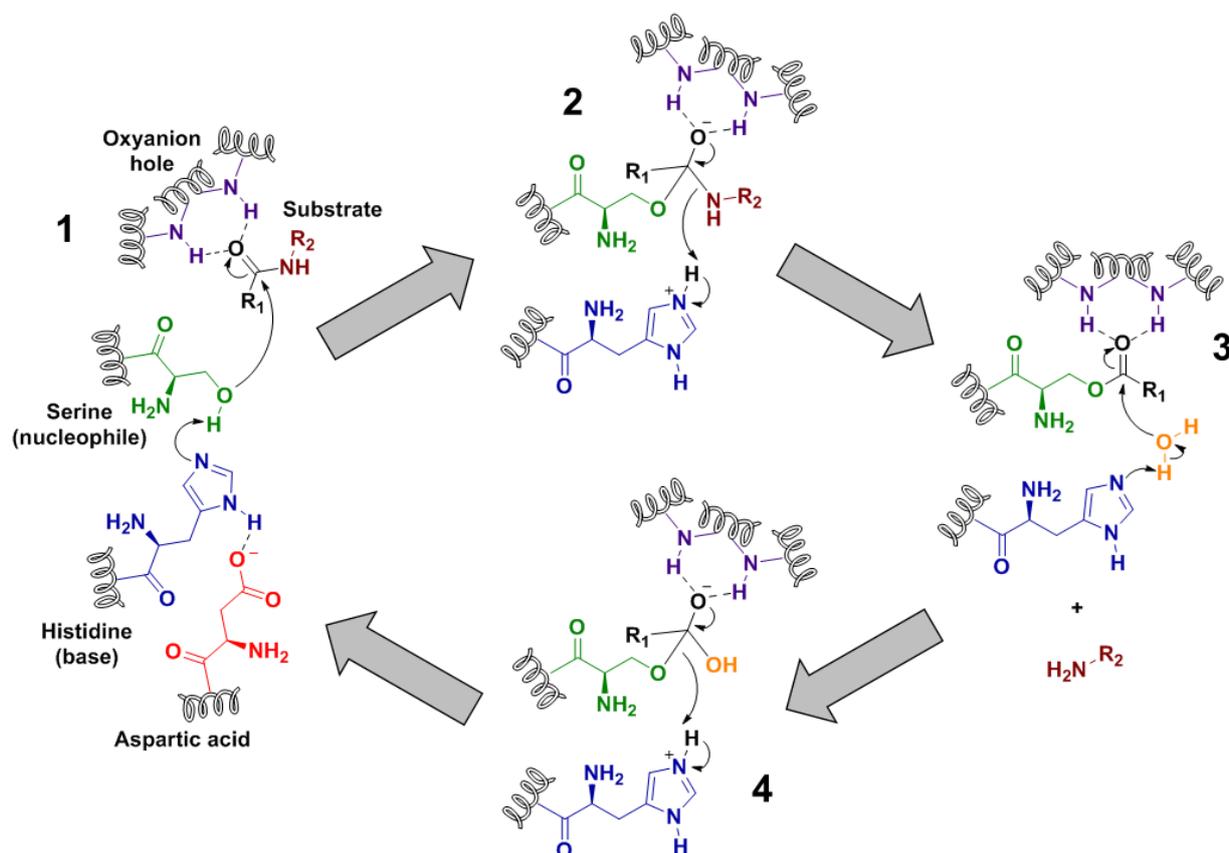


Figure 7.3: Example mechanism of catalytic triad undergoing covalent catalysis with serine as nucleophile (green), histidine as base (blue) and aspartic acid as acid (red).

An example mechanism for covalent catalysis by a catalytic triad is shown in **Figure 7.3**. As can be seen, triads are inter-dependant and perform in concert with other residues to achieve nucleophilic catalysis. A charge-relay system commonly found in serine proteases is depicted in the first step of **Figure 7.3.1**. The acidic triad member (red), typically an aspartate or glutamate residue, interacts with the histidine base (blue), thereby aligning and restricting its side-chain rotation and promoting stabilisation of the positive charge by polarisation. This allows for deprotonation and activation of the nucleophilic residue (green), exemplified by a serine residue. The three-dimensional structure of the enzyme assembles these residues in the precise position and orientation to affect this sequence.⁵

With its pKa lowered, the nucleophilic residue attacks the carbonyl of the substrate (maroon), leading to a tetrahedral intermediate, shown in **Figure 7.3.2**. The negative charge of this transition state is stabilised by an oxyanion hole (purple), a backbone pocket typically consisting of amides and other positively charged residues. This stabilisation lowers the activation energy required and promotes the catalysis.⁶ This transition state collapses back to the carbonyl and ejects the first half of the substrate (maroon), leaving the second half covalently bound to the enzyme as an acyl-enzyme intermediate shown in **Figure 7.3.3**. This is often promoted by proton transfer from the base member.

Enzymes containing a catalytic triad may catalyse one of two reaction types, which is decided in the second stage of catalysis. Nucleophilic attack of the acyl-enzyme intermediate by water (orange), as depicted in **Figure 7.3.3**, results in splitting of the original substrate and is characteristic of hydrolases. In contrary, nucleophilic attack by an additional substrate on the acyl-enzyme intermediate leads to transfer of this portion to this secondary substrate, typified by transferase enzymes.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

In our example mechanism, nucleophilic attack by water (orange) affects formation of a new tetrahedral intermediate (**Figure 7.3.4**) which is resolved by ejection of the original serine nucleophile (green), releasing the second product and regenerating the catalytic triad to allow for further catalysis.

Owing to the activated nature of these catalytic serine residues, numerous studies of their covalent targeting through probes and inhibitors can be found in the literature.⁷ However, examples for the successful targeting of their non-catalytic counterparts may also be found. Instances of serine-targeting electrophiles for both the catalytic and non-catalytic serine residues will be discussed in the following section. Also, the design and implementation of a covalent serine-targeting strategy, built upon the principles of the catalytic triad machinery, will be discussed in **Section 7.3**.

7.2.2 Serine-targeting Covalent Inhibitors

7.2.2.1 Previously Discussed Electrophiles

Many of the functional groups found suitable for investigation in this study have already been introduced in the previous chapter and will therefore be discussed only briefly. Reactive electrophiles such as sulfonyl fluorides have been shown to modify both catalytic and non-catalytic serine residues, particularly in serine proteases, and act as selective probes in a variety of enzymes.^{8,9} Nitriles are known to form strong hydrogen bond interactions as acceptors with serine amino acids and have been documented to form imidates with serine proteases.^{10,11} While archetypal Michael acceptors such as acrylamides are unreactive toward serine nucleophiles without activation by catalytic machinery, we nevertheless wished to investigate the reactivity of a vinyl sulfonamide. Intrigued by the additional electron-withdrawing capacity and hardier nature granted by the sulfonamide, we believed this electrophile could be compatible with the Ser797 residue.⁷ Lastly, α -haloketone warheads, similar to those attempted in the previous chapter, were amongst the first affinity probes utilised for serine proteases.⁷ In particular, chloromethyl ketones have found marked success as serine protease inhibitors, while their bromo- and iodomethyl analogues are typically more reactive, but less stable in aqueous solution, and are therefore used less frequently. In our opinion, inclusion of these functionalities and electrophiles in this research venture is warranted by their successful application in the literature.

7.2.2.2 Fluorophosphonates

Fluorophosphonates have been known as potent inhibitors of both serine proteases and esterases for almost a century.⁷ Prototype phosphonyl fluoride inhibitors included **sarin** and **soman**, shown in **Figure 7.4a** on the following page. Owing to their extreme reactivity with acetylcholinesterase, both agents are severely toxic and are regarded as weapons of mass destruction when considered as chemical warfare agents. The more stable and less reactive derivative, diisopropyl phosphonofluoridate (**DFP**) (**Figure 7.4a**), irreversibly inhibits many serine proteases and has found use in ophthalmology in the treatment of chronic glaucoma and as an irreversible anti-cholinesterase inhibitor.⁷ Similarly, methoxy arachidonyl fluorophosphonate (**MAFP**), inhibits nearly all serine hydrolases and proteases, displaying marked potency in the low nanomolar IC_{50} range in the inhibition of phospholipase A2 and fatty acid amide hydrolase. Although the above-mentioned substances lack resemblance to a peptide substrate, and as such retain no selectivity toward a particular serine protease or hydrolase, these agents have found use as diagnostic tools for the identification and classification of active site serine residues in these enzymes.⁷

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

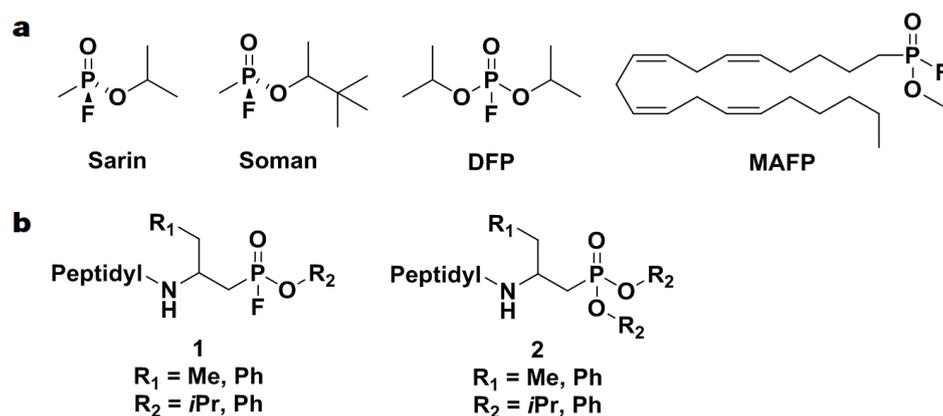


Figure 7.4: a) Examples of non-selective phosphonyl fluoride inhibitors of serine proteases and hydrolases, and b) Peptidyl substrate phosphonyl fluoride and phosphonate inhibitors able to covalently modify serine proteases.

To overcome this issue of selectivity, several peptide phosphonyl fluoride inhibitors, derived from and mimicking the serine peptide substrate, have been reported.¹² The general structure of these compounds can be seen as **1** in **Figure 7.4b** and these inhibitors have been shown to exhibit good selectivity towards their target serine proteases, especially that of chromotrypsin and elastase.⁷ The chemical stability and reactivity of these organophosphorus compounds in inhibiting serine proteases irreversibly is dictated by the electrophilicity of the central phosphorus atom. Regarding phosphonyl fluorides, the phosphorous being adjacent to an electronegative fluorine atom leads to high reactivity and low stability. This means that while effective as serine protease inhibitors, these agents are often unstable in aqueous conditions and undergo nonenzymatic hydrolysis.¹²

Compromising between reactivity and chemical stability, Oleksyszyn and Powers. developed the peptidyl *R*-aminoalkylphosphonates **2**, shown in **Figure 7.4b**, which were able to irreversibly inhibit serine proteases.¹³ While similar to the peptide phosphonyl fluorides **1**, the adjoining phenoxy groups are less electronegative than their fluorine counterparts. This engenders sufficient electrophilicity at the phosphorous atom to allow for reaction by the active site serine, while remaining hydrolytically stable.¹³ Due to the perceived weak reactivity of the Ser797 residue, we wished to explore the inclusion of highly reactive fluorophosphonate and alkylphosphonate electrophiles in our attempt to covalently modify the triple mutant serine residue.

7.2.2.3 Boronic Acids

Heretofore, the understanding and use of boron in therapeutics has been minimal relative to carbon, nitrogen, oxygen and sulphur.¹⁴ Boron has atomic number 5, with three electrons in its outer shell, and is considered a metalloid. The elements unique electronic configuration and ability to form covalent bonds with carbon has led to its wide use in synthetic chemistry.¹⁵ However, recent progress in the field of boron chemistry has expanded its application as synthetic reagent to use in the development of drug candidates.¹⁶ Most trivalent boron reagents are electrophilic, with the sp^2 hybridized trigonal boron atom able accept an electron pair from a nucleophile, leading to adoption of a sp^3 tetrahedral configuration and thereby satisfying the octet rule (**Figure 7.5a**). It is this ease of conversion from trigonal planar to tetrahedral that forms the basis of covalent inhibition of serine, threonine and lysine residues by boronic acid enzyme inhibitors.¹⁷ This is coupled with an inherent affinity to form strong reversible interactions with amine and hydroxyl functionalities in biological systems.¹⁴

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

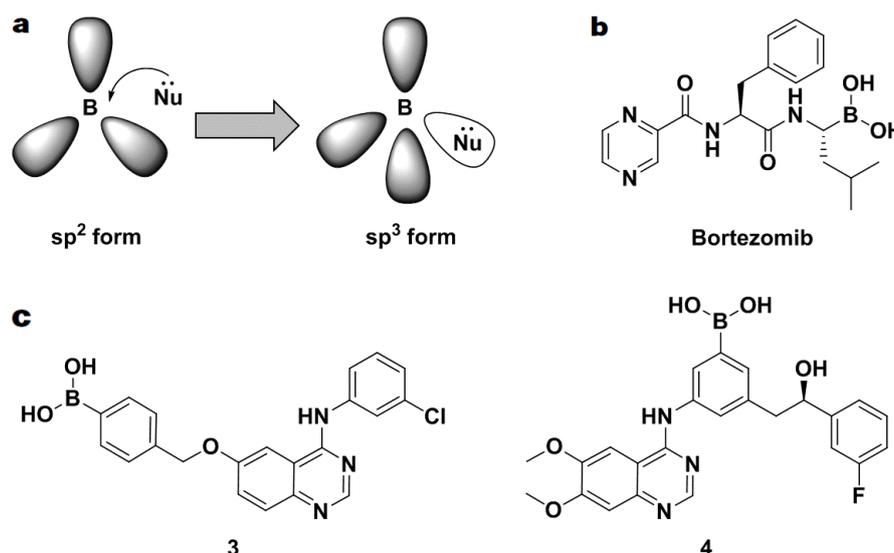


Figure 7.5: a) Configurational modifications of boron, b) Structure of the FDA approved inhibitor Bortezomib and c) Structure of boronic acid containing EGFR inhibitors **3** and **4**.

The pioneering exploration of boron's physical, chemical and biological properties by medicinal chemists has led to the representation of a new class of boron-based drugs. **Bortezomib**, shown in **Figure 7.5b**, received FDA approval in 2003 as the first proteasome inhibitor used for treatment of multiple myeloma and non-Hodgkins lymphoma and remains the only drug in clinical use with boron as an active element.¹⁸ The success of this discovery has motivated researchers to further explore boron chemistry, resulting in several drug candidates with anticancer, antiviral, antibacterial and other disease-specific activities reaching various phases of clinical trials.¹⁶

Implementation of boron chemistry has even extended into the development of EGFR inhibitors. As shown in **Figure 7.5c**, the boron-conjugated 4-anilinoquinazoline compound **3** was shown to effectively inhibit EGFR in a prolonged manner.¹⁹ Following mechanical docking simulations and time dependant *in-vitro* incubation assays, the prolonged inhibition was believed to result from irreversible inhibition of the enzyme, occurring through formation of a B-O bond with the Asp800 residue. Furthermore, separate studies have led to the development of compound **4**, showing single digit nanomolar efficacy against EGFR, while maintaining a good selectivity profile.²⁰

Due to the characteristic preference of boron for hard oxygen over soft sulphur nucleophiles, boronic acids have been utilised in the development of selective inhibitors of serine proteases.¹⁷ This is desirable in the context of our research objective and validates the inclusion of boronic acids in our arsenal of Ser797 targeting electrophiles.

7.3 Objectives and Target Synthetic Library

With the synthetic groundwork at our disposal and the desirable characteristics thereof discussed in **Section 6.2.2**, we chose to retain osimertinib as our heterocyclic scaffold for the synthesis of our target synthetic library. Similar to our lysine-targeting library, the design elements of our serine-targeting library may be found in **Figure 7.6** on the following page. Negating the need for molecular modelling, derivatisation with the selected serine-targeting electrophiles would occur at the aniline fragment which normally houses the acrylamide Michael acceptor in osimertinib. Incorporation at this position would place the warhead in an optimal orientation and proximal location to interact and potentially covalently modify the Ser797 mutant residue.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

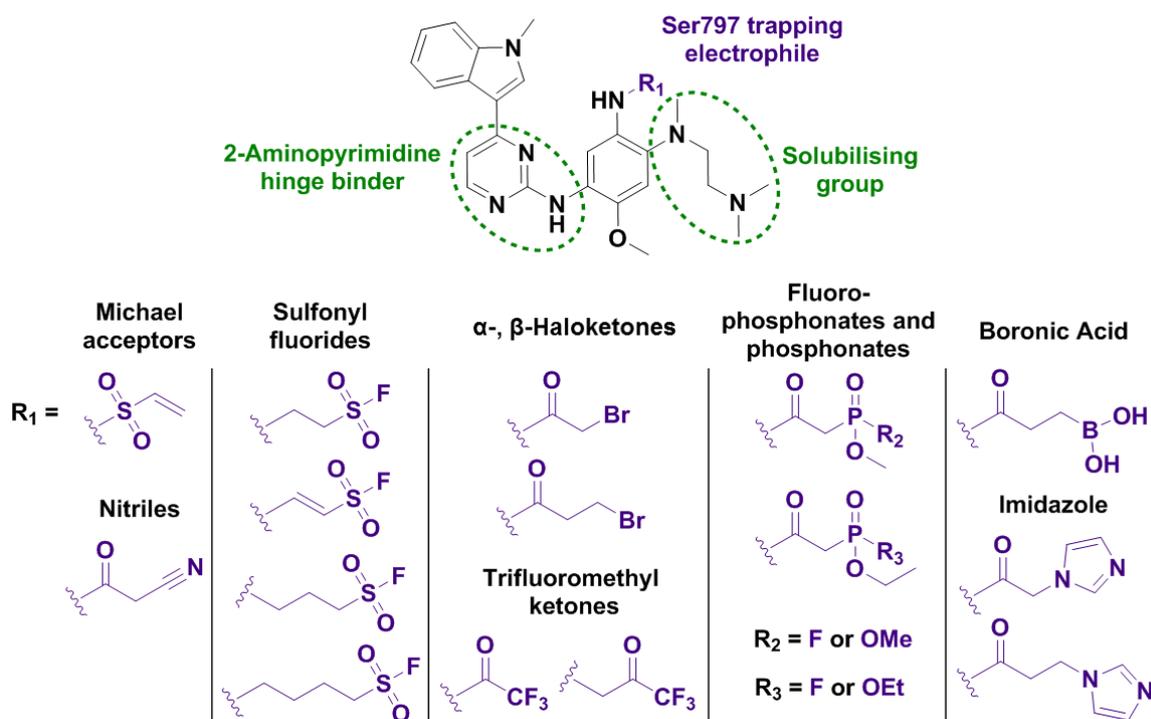


Figure 7.6: Target synthetic library for osimertinib-derived Ser797 targeting inhibitors.

Our envisaged electrophilic warheads can be seen above (purple, **Figure 7.6**) and include those discussed in previous instances such as ethenesulfonamide, cyanoacetamide, sulfonyl fluorides, α - and β -haloketones, fluorophosphonates and boronic acids. While not discussed in the same detail, we wished to investigate the potential interactions of the trifluoromethyl ketone functionality with the mutated Ser797 residue. Serine proteases have been shown to interact with trifluoromethyl ketones by forming stable hemiacetals.²¹ This exemplifies the covalent, albeit reversible, inhibitory effects that trifluoromethyl ketones may have on certain enzymes and their potential for inclusion in our serine-targeting library. Lastly, we wished to utilise the imidazole moiety to mimic the histidine catalytic triad base (**Figure 7.3**) and potentially establish favourable reversible interactions with the Ser797 residue.

Special consideration was given to the unsaturated sulfonyl fluoride functional group which is shown in **Figure 7.7a** on the next page. This specific electrophile can be seen as a “dual warhead,” as it contains both a Michael acceptor and a reactive sulfonyl centre adjacent to an electronegative fluorine atom which serves as a leaving group. This piqued our interest, as the two electrophilic centres are inherently selective in nature by their preference for either hard or soft nucleophiles. This infers that the Michael acceptor portion would facilitate covalent bond formation with the soft thiol nucleophile of a cysteine residue, while the sulfonyl fluoride fragment would be far more disposed to nucleophilic attack by a hard serine hydroxyl group (**Figure 7.7a**). If in concordance, this warhead could be capable of covalently modifying both the cysteine containing T790M double mutant and the mutated serine residue of the C797S triple mutant.

In keeping with the use of dual warheads, we devised a strategy built upon the previously discussed catalytic triad of serine proteases (**Figure 7.3**). The electrophiles shown in **Figure 7.7b** contain two functional groups, namely a cyano- or imidazole-acetamide fragment and a sulfonyl fluoride moiety.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

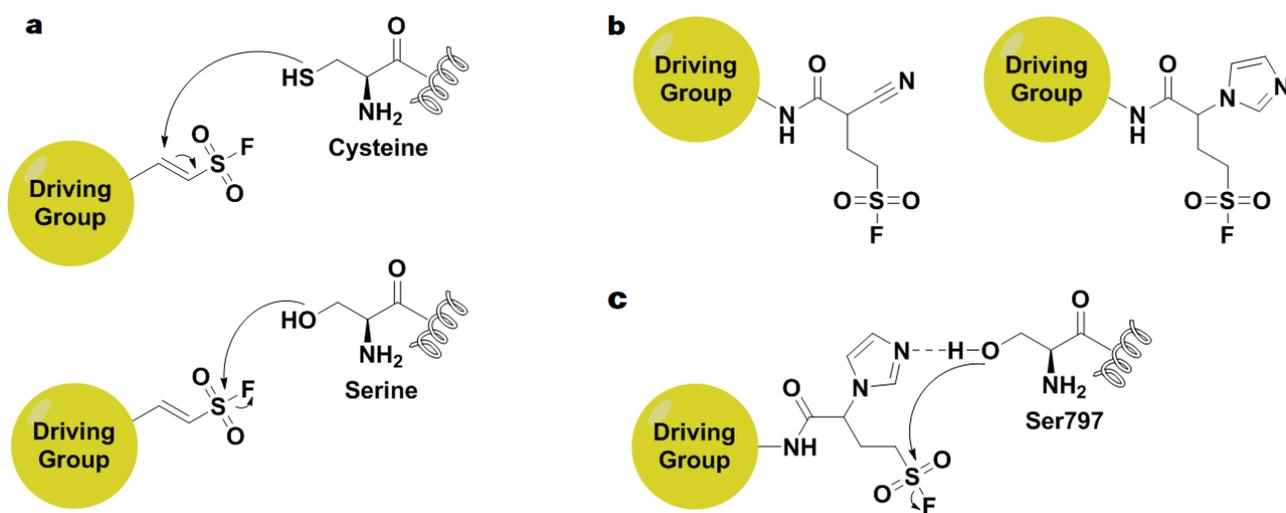


Figure 7.7: a) Dual warhead able to potentially covalently modify both cysteine and serine, b) Envisaged dual warheads and c) Mechanism of activation based on the catalytic triad of serine proteases.

Simulating the role of the histidine base in activating the serine residue of the catalytic triad, we envisaged the imidazole (or cyano) moiety interacting with Ser797 through hydrogen bonding (**Figure 7.7c**). With the electron withdrawing effects of the adjacent amide emulating the acid part of the catalytic triad in activating the base, these hydrogen bond interactions could in turn activate the Ser797 residue, potentially creating a more nucleophilic hydroxyl group. In conjunction with this, these interactions would provide a proximal location and orientation for the second part of the dual warhead, the electrophilic sulfonyl fluoride, to undergo nucleophilic attack from the activated Ser797. Intrigued by this prospect, we aimed at synthesising the cyano- and imidazole-containing dual warheads to explore the potential of our hypothesis, as well as the other compounds pertaining to our Ser797-targeting library.

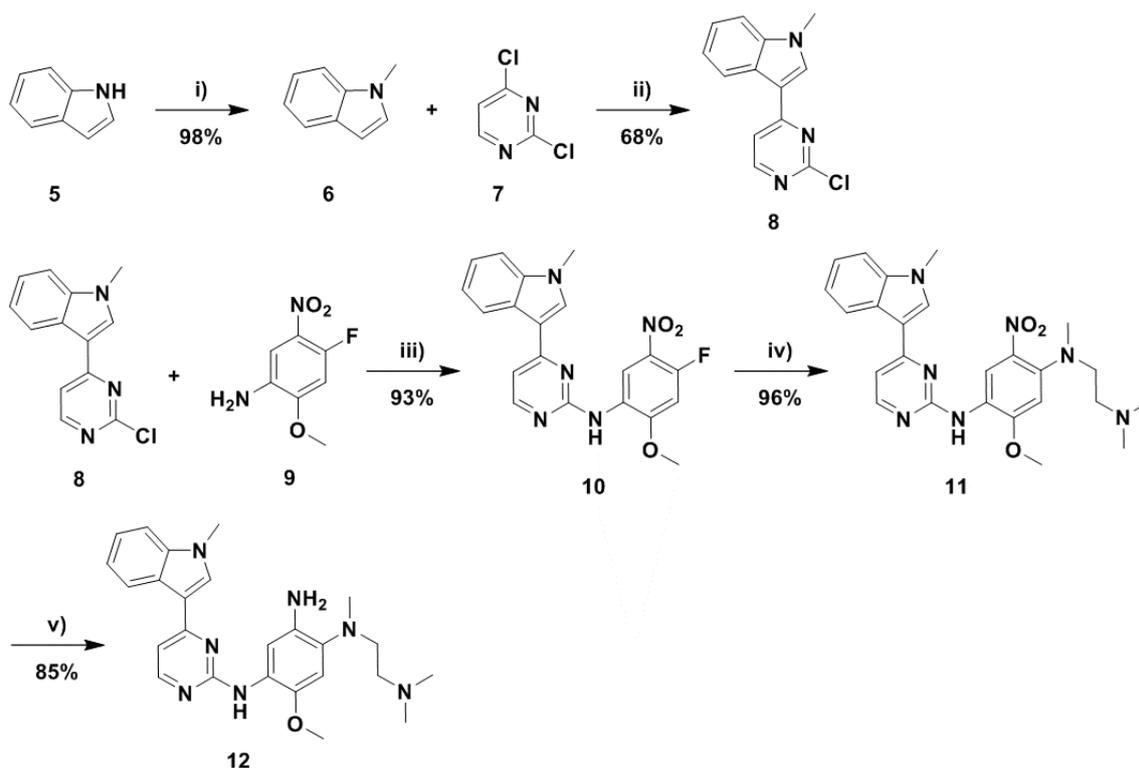
7.4 Library Synthesis

7.4.1 Osimertinib-derived Scaffold Synthesis

Owing to the incredible success of osimertinib in the clinical treatment of lung cancer, a wealth of approaches towards its synthesis can be found in the literature.^{22, 23} This allowed for the effective and large-scale synthesis of the osimertinib driving group scaffold, shown as compound **12** in **Scheme 7.1** on the following page. Starting from compound **5**, methylation of the indole nitrogen was carried out through deprotonation using sodium hydride as base, followed by dropwise treatment with methyl iodide.²⁴ Due to its exothermic nature, the reaction was carried out at $-10\text{ }^{\circ}\text{C}$ (acetone/ice), affording compound **6** in near quantitative yield. Coupling of the synthesised fragment **6** and the commercially available **7** was next accomplished using AlCl_3 as Lewis acid in DME. Use of this solvent allowed for a higher reaction temperature of $85\text{ }^{\circ}\text{C}$, which afforded the indole pyrimidine fragment **8** in 68% yield.²³

The remaining steps of the synthesis towards compound **12** was carried out in a linear fashion, similar to the route developed in the previous chapter. Furthermore, compound **9** was synthesised according to the procedure outlined in the previous chapter (**Section 6.4.1, 21**) and will therefore not be discussed.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

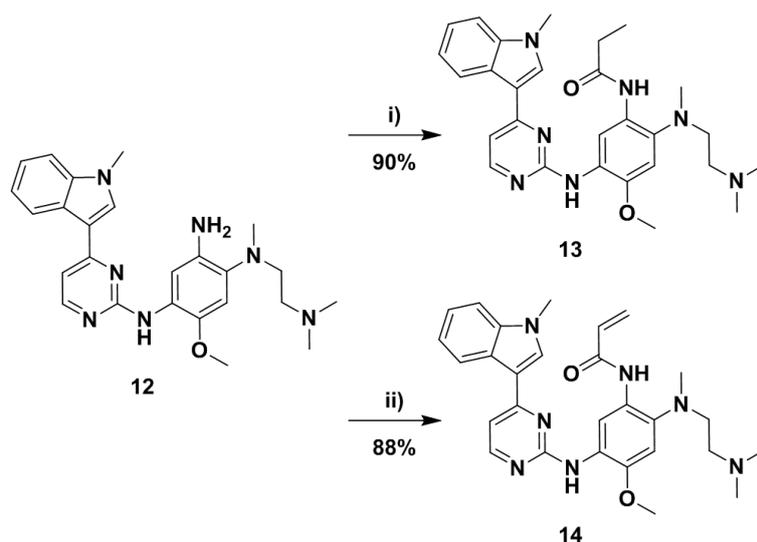


Scheme 7.1: Synthesis of the osimertinib driving group scaffold **12**. Reagents and conditions: i) NaH (2 equiv.), MeI (1.2 equiv.), THF, $-10\text{ }^{\circ}\text{C}$ – rt, 3 h; ii) **5** (1 equiv.), **7** (1.1 equiv.), AlCl_3 (1 equiv.), DME, $0\text{ }^{\circ}\text{C}$ – $60\text{ }^{\circ}\text{C}$, 12 h; iii) $\text{TsOH}\cdot\text{H}_2\text{O}$ (1.2 equiv.), 2-pentanol, $105\text{ }^{\circ}\text{C}$, 4 h; iv) N^1,N^1,N^2 -trimethylethane-1,2-diamine (1.25 equiv.), DIPEA (1.05 equiv.), DMA, $60\text{ }^{\circ}\text{C}$, 4 h; v) Fe (6 equiv.), NH_4Cl (1 equiv.), EtOH/ H_2O (3:1), $85\text{ }^{\circ}\text{C}$, 4 h.^{23,24}

Coupling between compound **8** and **9** was achieved using the previously optimised reaction conditions, employing *p*-toluenesulfonic acid monohydrate in 2-pentanol at $105\text{ }^{\circ}\text{C}$. Following collection of the precipitate by filtration and a simple work-up procedure, the coupled fragment **10** was afforded in the excellent yield of 93% with no further purification required. Installation of the solubilising group was carried out in the subsequent nucleophilic aromatic substitution reaction, furnishing compound **11** as a bright red solid. Lastly, reduction of the aromatic nitro group was executed using iron and ammonium chloride in a 3:1 mixture of ethanol/water while heating under reflux. This afforded compound **12** in 85% yield, with the obtained ^1H , ^{13}C and MS data comparing well with the reported literature values, and concluded the synthesis of our osimertinib driving group scaffold.^{22,23}

7.4.2 Model Compounds

To allow for comparative analysis with our library of serine-targeting compounds in the biochemical and cellular evaluation, we envisaged synthesis of both the non-covalent propionamide analogue **13** and osimertinib (**14**) as model compounds. As illustrated in **Scheme 7.2**, both compounds were synthesised using acylating agent, potassium carbonate as base and acetone as solvent. However, yields for compound **14** were found to improve dramatically through implementation of low temperatures generated by an acetonitrile and dry ice bath, thereby avoiding polymerisation of the acryloyl chloride and side product formation. This furnished both **13** and **14** in excellent yields of 90 and 88% respectively which were fully characterised using IR, MS and NMR spectroscopy.

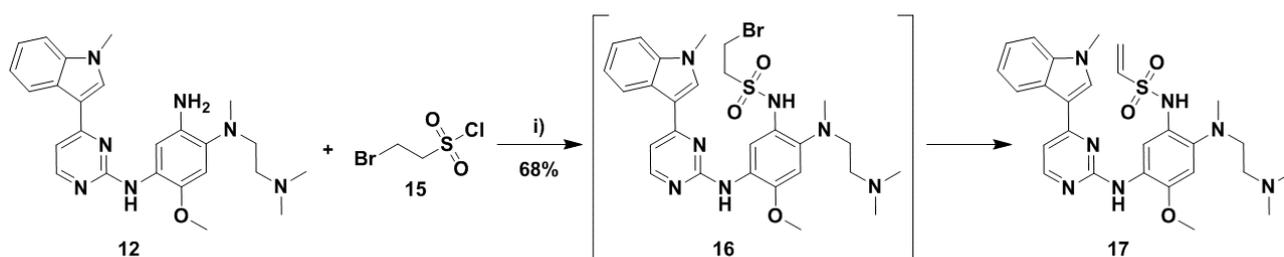


Scheme 7.2: Synthesis of compounds **13** and **14**. Reagents and conditions: i) propionyl chloride (1.3 equiv.), K_2CO_3 (1.5 equiv.), acetone, $-10\text{ }^\circ\text{C}$ – rt, 2 h; ii) acryloyl chloride (1.1 equiv.), K_2CO_3 (1.5 equiv.), acetone, $-40\text{ }^\circ\text{C}$ – $-20\text{ }^\circ\text{C}$, 2 h.

7.4.3 Electrophile Synthesis and Derivatisation

7.4.3.1 Ethenesulfonamide

With the effective synthesis of our osimertinib driving group scaffold **12**, we set our sights on aniline derivatization with the various designated serine-targeting electrophiles. Our synthetic pursuit began within the Michael acceptor category, with the singular ethanesulfonamide **17** shown in **Scheme 7.3** below. Due to the more nucleophilic nature of aniline **12**, we did not require the problematic use of strong base and harsh reaction conditions which led to the unsuccessful isolation of the indole analogue in the previous chapter. Initially, reaction of **12** with the previously synthesised **15**, using one equivalent of DIPEA in DCM, led to the formation of two prominent spots when visualised by TLC analysis. Interestingly, upon exposure to a dilute basic solution during workup, only one of the two spots remained. Purification of this crude product using column chromatography was followed by NMR spectroscopic analysis, confirming the successful isolation of compound **17**. We gathered that spontaneous elimination of intermediate **16** had occurred during the workup, easily ascribed to the good bromine leaving group and the highly acidic nature of the proton *alpha* to the electron-withdrawing sulfonamide which both facilitate elimination. Indeed, subsequent reactions using an excess of base resulted in the development of one spot when visualised by TLC analysis and the sole formation of **17**. This strategy furnished ethenesulfonamide **17** in a good yield of 68%, with discussion of the ^1H NMR spectroscopic characterisation to follow.



Scheme 7.3: Synthesis of ethenesulfonamide **17**. Reagents and conditions: i) **15** (1.3 equiv.), DIPEA (3 equiv.), DCM, $-40\text{ }^\circ\text{C}$ – $-20\text{ }^\circ\text{C}$, 2 h.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

The ^1H NMR spectra of compound **17** can be found below in **Figure 7.8**, giving structural confirmation for the synthesis of the ethenesulfonamide. Examination of the most upfield region reveals the presence of a triplet at δ 2.27 ppm, with the peak mirrored by another triplet at δ 2.87 ppm, both integrating for 2H and with identical J-coupling values. These two triplets were assigned to the ethylene bridge (dark blue) of the solubilising group. The two large singlets in between constitute the remainder of the solubilising group, with the peak at δ 2.35 ppm integrating for 6H ascribed to the dimethylamine (green) element and the signal integrating for 3H at δ 2.73 ppm correlating to the methylamine (light blue) of the solubilising group. Moving downfield, two characteristic methyl signals, both integrating for 3H, could be found at δ 3.90 and 3.97 ppm. Based on spectra of previously synthesised building blocks, these peaks were assigned to the indole methyl (orange) and the aryl methoxy (yellow) respectively. Indication of the ethenesulfonamide moiety (red) was found by the three signals between δ 5.70 and 6.65 ppm. These signals comprised of two doublets at δ 5.74 and 6.16 ppm ($J = 10.0$ Hz and 16.7 Hz respectively), and a doublet of doublets δ 6.58 ppm ($J = 16.7$ Hz), characteristic of vinyl sulfone protons. Assessment of the aromatic region of the ^1H NMR spectrum presented the correct integration amount for compound **17**, with all protons accounted for and peak splitting patterns within expectations. This left us with two singlets at the farthest region downfield of the spectrum, corresponding to the remaining two amine protons within our structure. Working with NMR spectra of precursory structures and the inherent electron-withdrawing nature of the sulfonamide, the broader singlet at δ 8.70 ppm was assigned to bridging aniline (purple) and the sharper singlet furthest downfield at δ 9.02 ppm to the sulfonamide proton (brown). While not displayed for interpretation, all carbons were correctly accounted for in the corresponding ^{13}C spectra. Furthermore, use of HRMS analysis presented a $[\text{M}+\text{H}]^+$ ion of 536.2450, with a calculated mass of 536.2444, providing final confirmation for the successful synthesis of compound **17**.

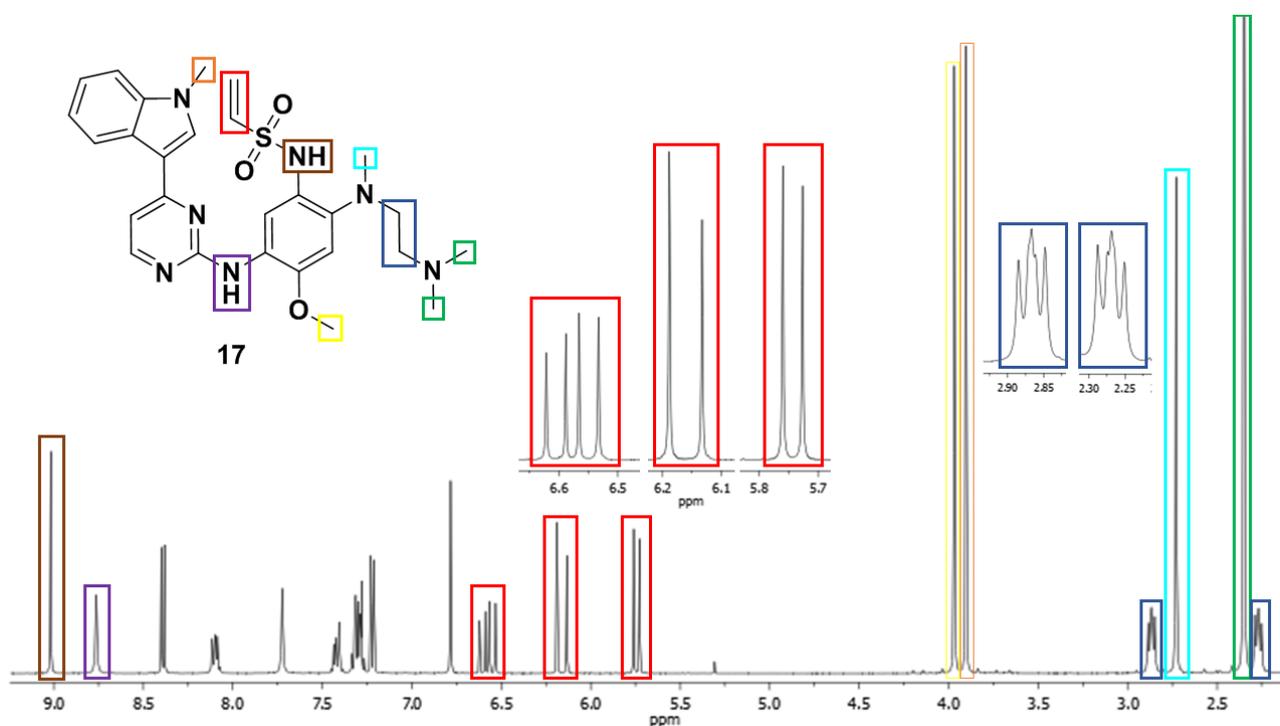
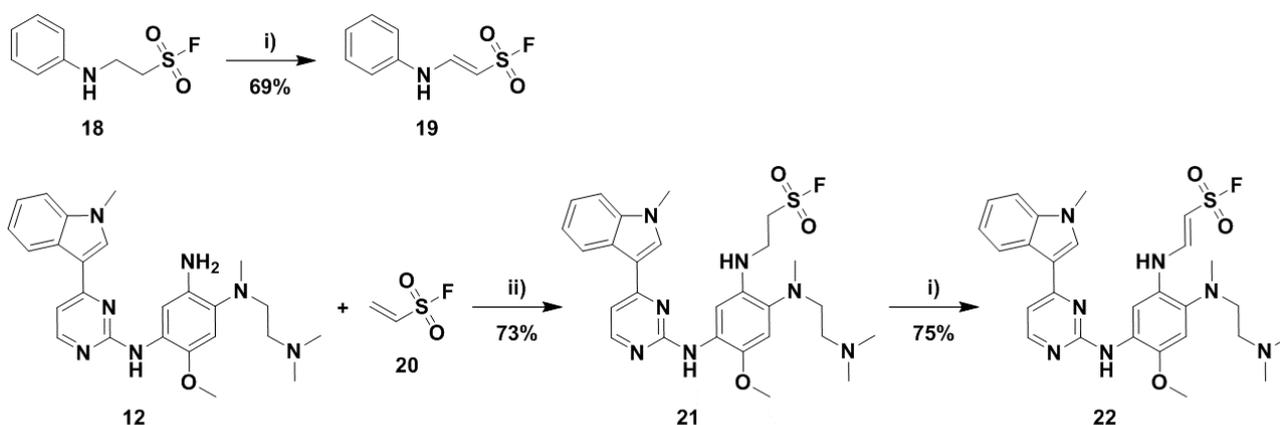


Figure 7.8: ^1H NMR spectra of compound **17**.

7.4.3.2 Sulfonyl Fluorides



Scheme 7.4: Synthesis of model dual warhead **19** and the osimertinib derived sulfonyl fluoride **21** and dual warhead **22**.

Reagents and conditions: i) MnO₂ (15 equiv.), CHCl₃, rt, 8 h; ii) ESF (1.05 equiv.), DMF, -10 °C – rt, 4 h.

In the previous chapter, compound **18**, shown in **Scheme 7.4** above, was synthesised as a model to perform comparative spectroscopic analysis with the sulfonyl fluoride molecules synthesised using ¹H and ¹⁹F NMR spectroscopy. Wishing to repeat the successful application of this principal, we synthesised the dual warhead compound **19** as a model representation (**Scheme 7.4**). Use of excess manganese dioxide in chloroform facilitated smooth oxidation to the alkene at room temperature, affording **19** in a yield of 69% without optimization.²⁵

Utilising similar conditions to afford sulfonyl fluorides in the previous chapter, scaffold **12** was dissolved in DMF and treated dropwise with ESF (**20**) at low temperature (**Scheme 7.4**). As before, the Michael addition was found to proceed effectively without the use of base. Following basic work-up and purification of the crude product using flash column chromatography, compound **21** was afforded as a yellow solid in good yield. This allowed us to attempt conversion of **21** into **22** using the same oxidation reaction parameters developed for compound **19**. To our delight, the osimertinib-derived dual warhead **22** was furnished in an improved yield of 75% when compared with the model system. To the best of our knowledge, besides oxidation of **18** to the model compound **19**, no other application of this reaction sequence exists within the literature, highlighting the obscurity and potential for use of this class of warhead.

The most discernible substantiation for the successful synthesis of compounds **21** and **22** can be found in the comparison of their respective ¹⁹F spectra with that of the model compounds **18** and **19**, illustrated in **Figure 7.9**. For the model compounds (**Figure 7.9a**), the saturated sulfonyl fluoride (red) **18** appears as a well-formed triplet at δ 56.7 ppm, coupling to the neighbouring protons of the ethyl bridge carbon. Following oxidation, the model dual warhead **19** fluoride peak (blue) shifts approximately δ 15.4 ppm downfield to arise at δ 72.1 ppm. This is anticipated, as formation of the conjugated alkene system would withdraw electron density from the sulfonyl system, thereby deshielding and causing a downfield shift in the signal. Furthermore, the signal multiplicity changes from that of a triplet to a doublet, in line with the loss of coupling to one proton in the unsaturated system.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

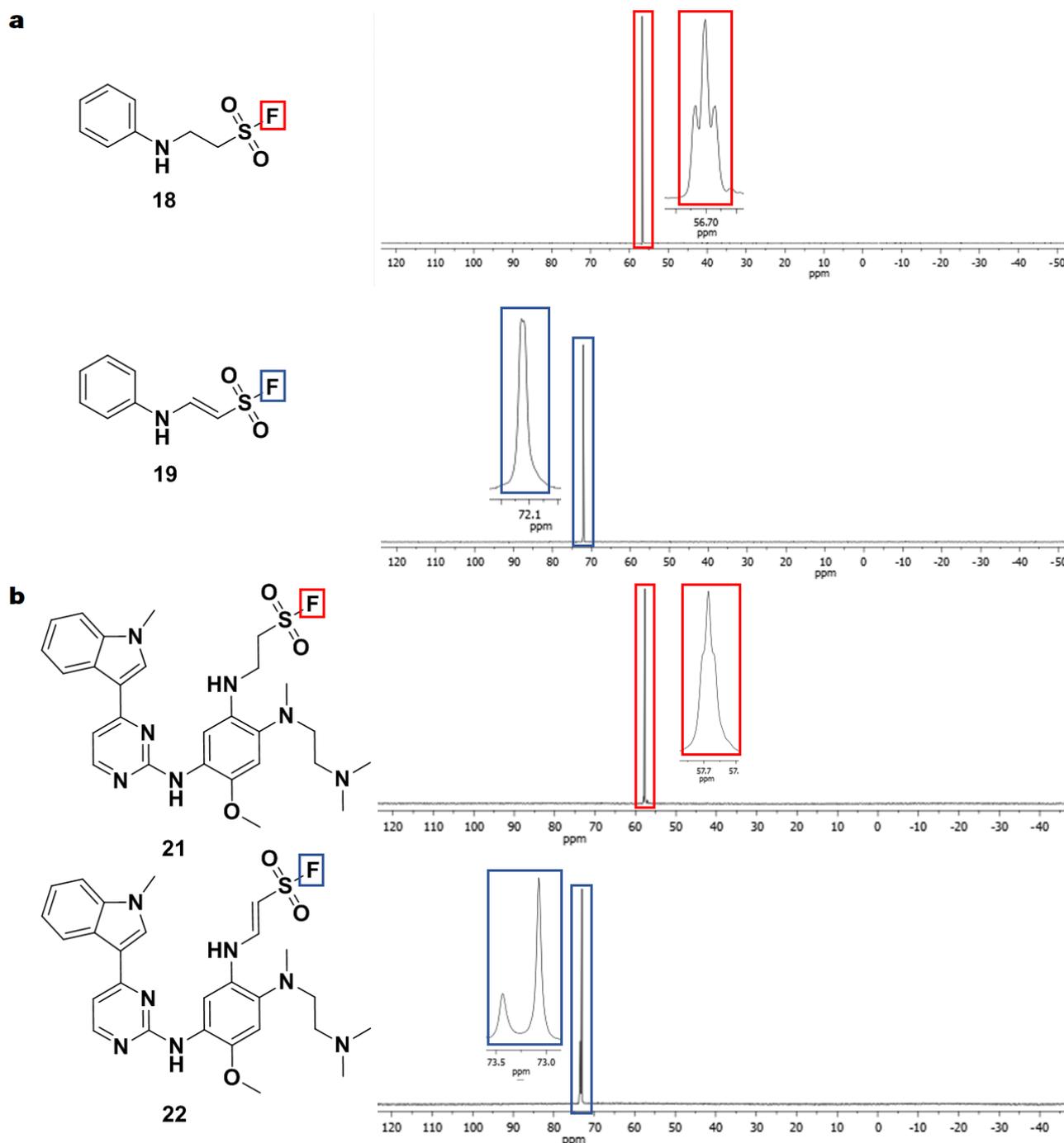
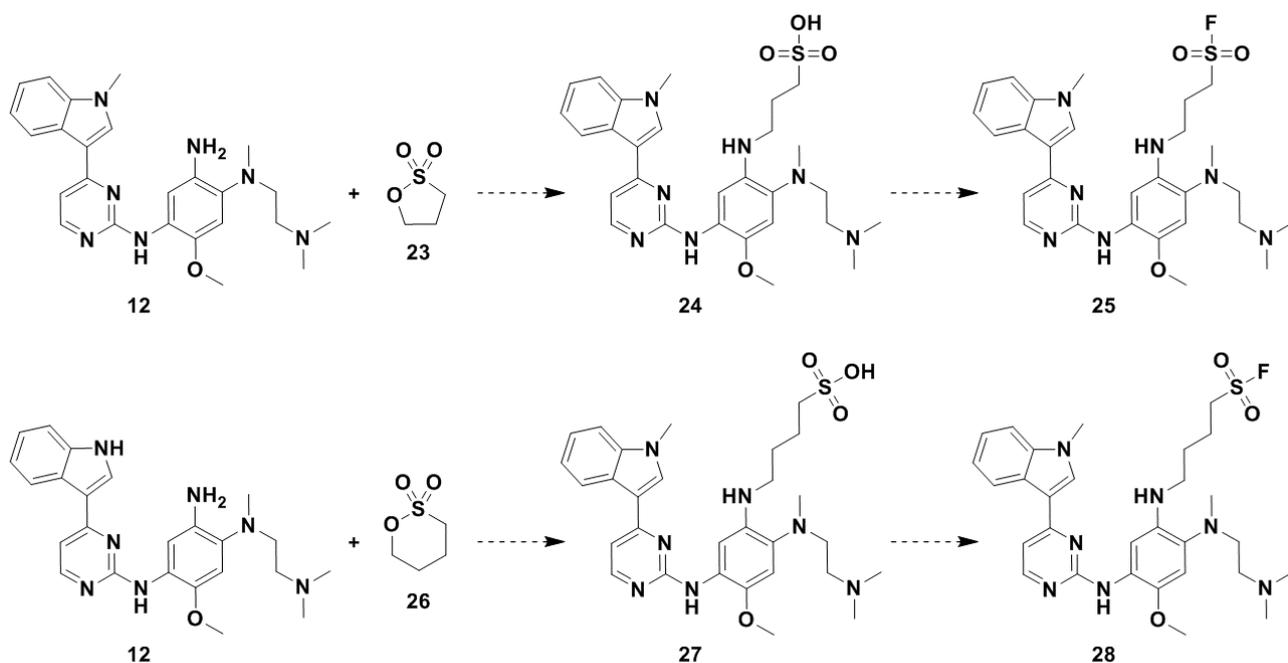


Figure 7.9: Comparison of the ^{19}F NMR spectra of a) model compounds **18** and **19** and b) final compounds **21** and **22**.

The model compounds spectroscopic features are mirrored in the ^{19}F NMR spectra of **21** and **22**, found above in **Figure 7.9b**. Parallel to its model counterpart, the fluorine signal of compound **21** appears at δ 57.7 ppm maintaining an albeit broadened triplet multiplicity. Analogous to the shift found in the model system, the signal of **22** moves downfield by roughly δ 15.5 ppm to emerge at δ 73.2 ppm, noticeably changing to the expected doublet. Further corroborating the successful synthesis of these compounds, all signals within the ^1H and ^{13}C NMR spectra were correctly accounted for. Final validation was provided by HRMS analysis, revealing a $[\text{M}+\text{H}]^+$ ion of 556.2494, with calculated mass of 556.2506 for **21** and a $[\text{M}+\text{H}]^+$ ion of 554.2352 for **22**, with a calculated mass of 554.2350.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



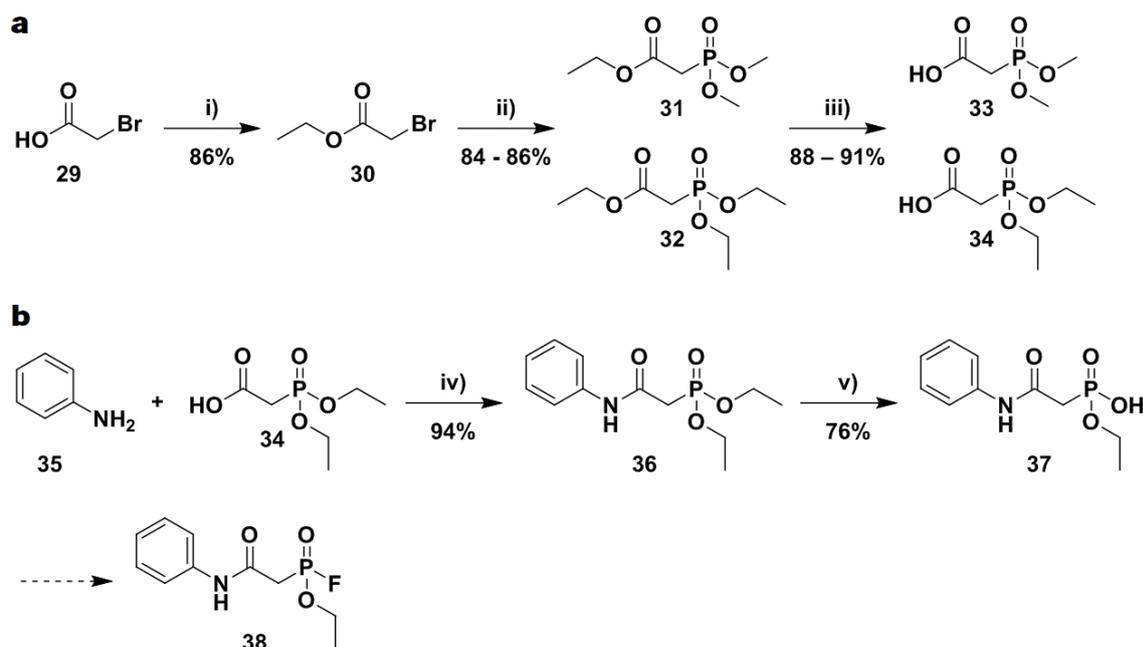
Scheme 7.5: Attempted synthesis of final compounds **25** and **28**.

The attempted synthesis of the propyl (**25**) and butyl (**28**) bridge sulfonyl fluoride counterparts is shown above in **Scheme 7.5**. Unfortunately, due to the same issues experienced in the synthesis of the previous chapter analogues, this synthesis was met with failure and we were unable to obtain these compounds. The major pitfalls associated with this failure surrounded the isolation and purification of the sulfonic acid intermediate compounds **24** and **27**. As these difficulties were outlined in detail at the end of **Section 6.4.2.3**, we chose not to discuss this synthesis further, thereby concluding the synthesis of our sulfonyl fluoride library of serine-targeting inhibitors.

7.4.3.3 Phosphonates and Fluorophosphonates

To furnish our phosphonate and fluorophosphonate based serine-targeting inhibitors, we envisaged the synthesis of the phosphonate fragments **33** and **34** found in **Scheme 7.6a**. Thereafter, subsequent coupling acid coupling, mono-hydrolysis and fluorination could afford the fluorophosphonate-containing compounds able to potentially covalently modify the mutant Ser797. Starting from the commercially available bromoacetic acid (**29**), esterification using sulfuric acid as catalyst, whilst stirring in ethanol heated under reflux, afforded the ethyl ester (**30**) in 86% yield.²⁶ This allowed us to undertake the Michaelis-Arbuzov reaction between alkyl halide **30** and a respective trivalent phosphorous ester to afford the corresponding alkyl phosphonate. Accordingly, trimethyl- or triethyl-phosphite were rapidly added to neat ethyl bromoacetate (**30**) at 80 °C, followed by swiftly ramping the temperature up to 130 °C, which allowed for distillation of the bromoethane by-product released during the formation of **31** and **32**.²⁷ Purification was achieved by kugelrohr vacuum distillation, with fractions boiling between 145 – 150 °C and 150 – 155 °C at ~ 20 mbar collected to afford phosphonate esters **31** and **32** respectively, as clear oils. Subsequent hydrolysis of the ethyl ester moiety for both **31** and **32** was carried out by treatment with an equimolar amount of sodium hydroxide, followed by acidification to pH = 1 with 1 M HCl and aqueous extraction. Finally, removal of the solvent *in vacuo* afforded both carboxylic acid phosphonate fragments **33** and **34** in excellent yield with no further purification required.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



Scheme 7.6: Synthesis of a) phosphonate electrophile fragments **33** and **34** and b) model compound **36** and **37**. Reagents and conditions: i) H_2SO_4 (cat.), EtOH, 85 °C, 24 h; ii) trialkylphosphite (1 equiv.), neat, 80 °C – 130 °C, 3 h; iii) NaOH (1 equiv.), 1 M HCl, H_2O , rt, 2 h; iv) **34** (1.85 equiv.), CDI (1.85 equiv.), DCM, 0 °C – rt, 12 h; v) MgCl_2 (1.2 equiv.), NaI (1.2 equiv.), MeCN, 60 °C, 12 h.

To gain a better understanding of the handling and reactivity of the desired fluorophosphonates, and the reaction sequence to be applied to driving group scaffold **12**, we undertook a preliminary model synthesis towards fluorophosphonate **38**, shown in **Scheme 7.6b**. Mono-hydrolysis of the phosphonate ester and isolation of the highly reactive fluorophosphonate was anticipated to be particularly problematic. Following the employment of numerous coupling reagents, it was found that treatment of excess carboxylic acid **34** and CDI with aniline (**35**) in DCM facilitated optimal amide bond formation. This provided **36** in the excellent yield of 94%, allowing us to tackle mono-hydrolysis of the ethyl ester phosphonate.

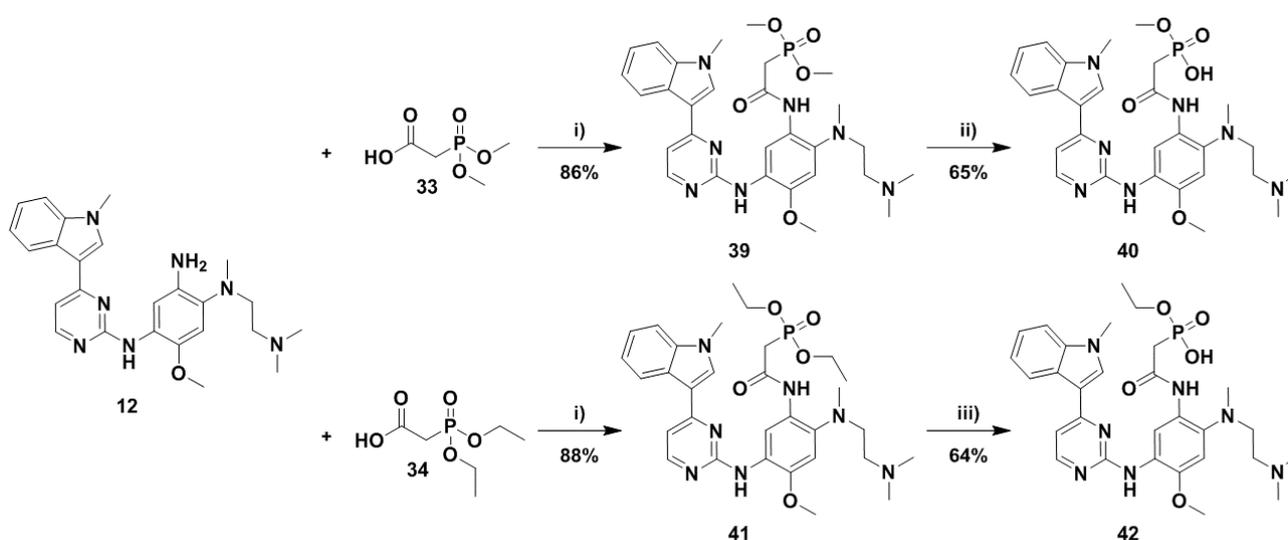
The selective mono-hydrolysis of **36** proved to be a challenging endeavour, with initial attempts being met with failure. Conventional hydrolysis methods utilising NaOH or KOH led to a loss of hydrolytic control, resulting in the formation of the di-hydrolysed phosphonic acid. We therefore turned towards investigation of Lewis acids with varying strength as potential hydrolytic reagents. Use of the weaker Lewis acids LiCl and LiBr in high boiling point solvents such as toluene, acetonitrile and N-methyl-2-pyrrolidone resulted in no reaction at varying temperatures.²⁸ By contrast, use of stronger Lewis acids such as AlCl_3 and TiCl_4 provoked di-hydrolysis and the formation of a several by-products, which was also the case when implementing use of the silylating agent TMSCl with various additives.²⁹ Following many attempts and much frustration, successful mono-hydrolysis was finally attained using the unfamiliar Lewis acid magnesium chloride with sodium iodide as additive.²⁹ Optimal reaction conditions were found to require acetonitrile as solvent at a temperature of 60 °C. This provided the mono-hydrolysed phosphonate **37** in good yield, with successful synthesis confirmed by HRMS and ^1H , ^{13}C and ^{31}P NMR spectroscopic analysis. Distinct shifts in the signals of the ^{31}P clearly evinced this and will be used for comparative analysis with the final compound analogues at a later stage.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Formation of fluorophosphonate **38** was attempted using the fluorinating reagent DAST at temperatures ranging from $-78\text{ }^{\circ}\text{C}$ – $0\text{ }^{\circ}\text{C}$ in both DCM and DMSO. While we believe that these efforts were successful, owing to the disappearance of starting material **37** and formation of a new spot when visualised by TLC analysis, the instability of **38** out of solution unfortunately impeded structural confirmation by spectroscopic analysis. It was due to this compound degradation that we were unable to isolate **38** and while disheartening, we still endeavoured to attempt the synthesis of the fluorophosphonate final compound counterparts.

The previously determined model compound reaction parameters were then adopted to our final compound synthesis, illustrated in **Scheme 7.7**. Using excess carboxylic acid and CDI in DCM, amide bond formation was successfully carried out between driving group scaffold **12** and **33** to afford the dimethyl phosphonate compound **39** in excellent yield. While no model reaction was carried out for the mono-hydrolysis of the dimethyl phosphonate ester moiety, we discovered a procedure by Gray and Smith in the literature which describes the selective cleavage of methyl phosphorous esters by *tert*-butylamine.³⁰ The researchers found that use of *tert*-butylamine as solvent, while heating under reflux, effected demethylation of methyl phosphonate esters exclusively. Furthermore, dimethyl phosphonates were selectively mono-hydrolysed due to the formation of a sterically bulky *tert*-butylamine salt intermediate, effectively blocking the reaction site and preventing di-hydrolysis. Gratifyingly, application of these conditions to **39** afforded the mono-hydrolysed phosphonate ester **40** in acceptable yield. Compound **40** was isolated by treatment of the crude product with HCl in dioxane, followed by purification by trituration.

For the ethyl phosphonate compounds, acid coupling between **12** and **34** was affected by the same reaction conditions described above to afford diethyl phosphonate **41** in 88% yield. Similarly, conditions derived from the model reaction sequence successfully facilitated the subsequent mono-hydrolysis of **41** to provide final compound **42** in a yield of 64%,



Scheme 7.7: Synthesis of phosphonate compounds **40** and **42**. Reagents and conditions: i) **33** or **34** (1.85 equiv.), CDI (1.85 equiv.), DCM, $0\text{ }^{\circ}\text{C}$ – rt, 4 h; ii) *tert*-butylamine, $55\text{ }^{\circ}\text{C}$, 16 h; iii) MgCl_2 (1.2 equiv.), NaI (1.2 equiv.), MeCN, $60\text{ }^{\circ}\text{C}$, 12 h.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

As was performed with the synthesised sulfonyl fluoride inhibitors, comparative analysis of the respective ^{31}P NMR spectra of model compounds **36** and **37** and final compounds **41** and **42** may be found below in **Figure 7.10**. All ^{31}P NMR spectroscopy experiments were self-run, with chemical shifts within spectra referenced to 85% H_3PO_4 in H_2O , which served as the internal standard. Correlation of this spectroscopic data provides verification of both the diethyl phosphonate and mono-hydrolysed analogue. As can be seen in **Figure 7.10a**, the phosphorous signal of **36** (red) appeared as a singlet at δ 23.4 ppm. Mono-hydrolysis of this model compound induces an upfield shift of δ 10.6 ppm, resulting in the signal of **37** appearing at δ 12.8 ppm. Furthermore, this peak is visibly broader than that of its predecessor, possibly due to interactions with the newly formed protic hydroxyl moiety.

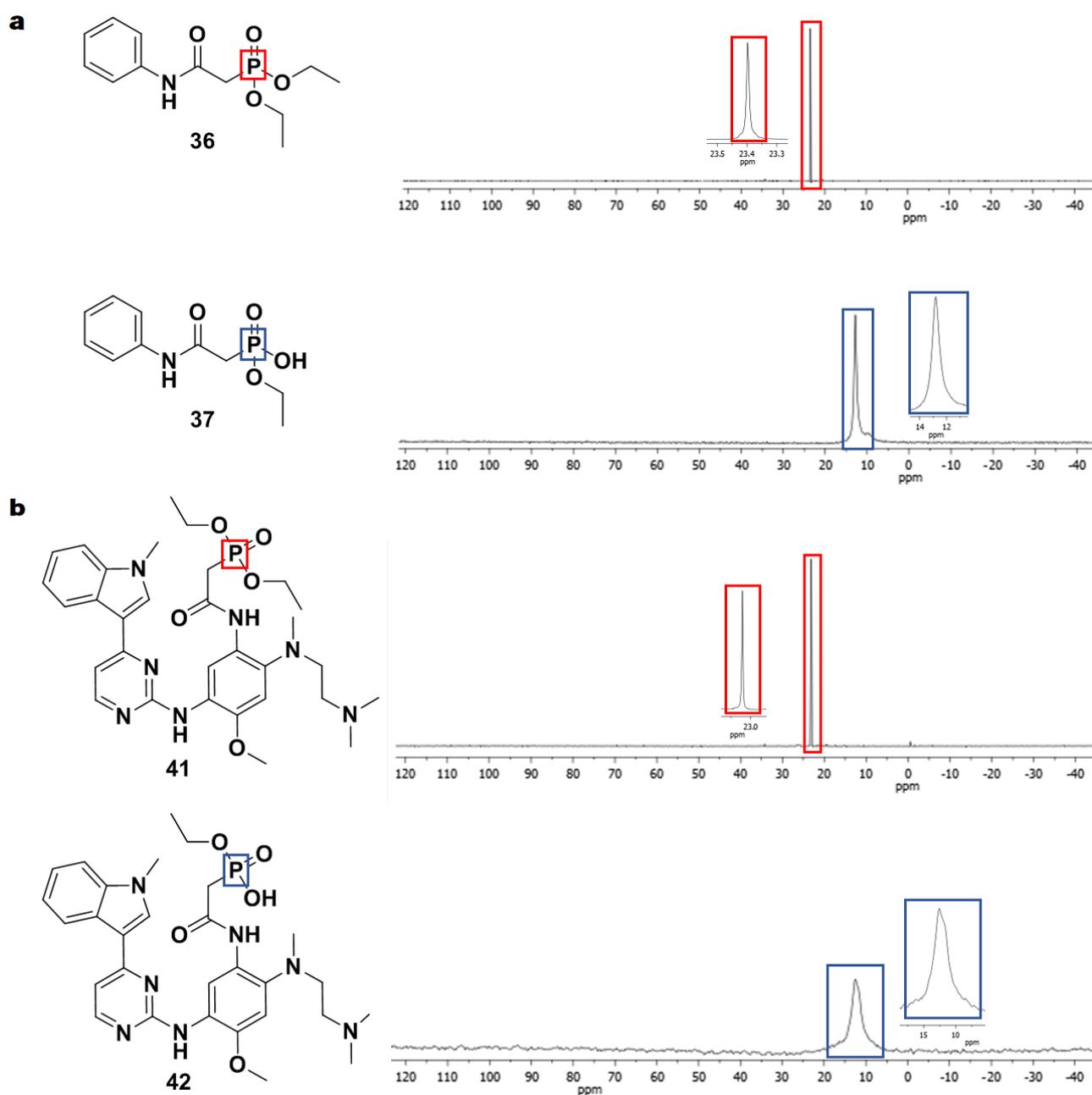


Figure 7.10: Comparison of the ^{31}P NMR spectra of a) model compounds **36** and **37** and b) final compounds **41** and **42**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Inspection of the final compounds **41** and **42** ^{31}P NMR spectra, shown in **Figure 7.10b**, reveals a near indistinguishable reflection of that seen for the model compounds. Comparable to its model equivalent, the phosphorous signal of **41** shows up at δ 23.1 ppm with a similar elongated peak shape. Following mono-hydrolysis, the anticipated peak shift for the phosphorous signal of compound **42** is identical to that of the model system, moving δ 10.6 ppm upfield to appear at δ 12.5 ppm. Additionally, the peak shape of **42** is observably broader, mirroring that of **37**. Beyond the use of ^{31}P NMR spectroscopy, all signals within the ^1H and ^{13}C NMR spectra were correctly accounted for, with the integration of the alkyl phosphonate signals correctly changing from a total of 10H to 5H. Lastly, HRMS analysis revealed a $[\text{M}+\text{H}]^+$ ion of 623.2920, with calculated mass of 623.2925 for **41** and a $[\text{M}+\text{H}]^+$ ion of 596.2769 for **42**, with a calculated mass of 596.2750. While not shown or directly discussed, the same considerations were evident in all analytical data retrieved for the methyl phosphonate compounds **39** and **40**, confirming the successful synthesis of all phosphonate final compounds.

The attempted synthesis of the fluorophosphonate analogues of these compounds was met the same pitfalls associated with product instability found in the model system. Sending the mono-hydrolysed phosphonate esters **40** and **42** to our colleagues at the lab of Prof. Daniel Rauh in Germany, we had hoped that the fluorination could be carried out and followed by immediate biochemical and cellular evaluation. Unfortunately, this plan did not come to fruition due to the highly reactive and instable nature of the fluorophosphonate group. This concluded the synthesis of our phosphorous containing serine-targeting inhibitors.

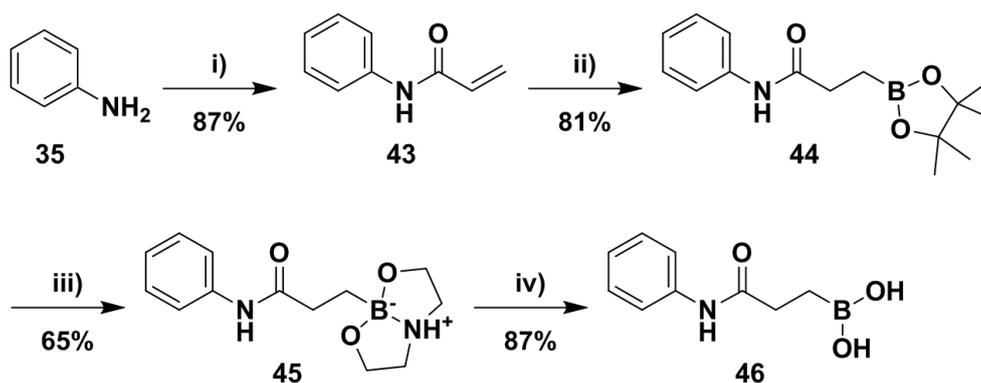
6.4.3.4 Boronic Esters and Acids

The introduction of a boronic ester and acid motif within our serine-targeting library would require a multi-step synthetic protocol. As such, we proposed the synthesis of the model boronic ester and acid compounds illustrated in **Scheme 7.8** on the following page. By undertaking this synthesis, we wished to reproduce the previous successes attained in model systems with respect to method development of optimal reaction conditions and comparative spectroscopic analysis.

Embarking on the model synthesis, the first step entailed acylation of aniline (**35**) with acryloyl chloride while cooling with an acetonitrile/dry ice bath, furnishing acrylamide **43** as a white solid in excellent yield. Building on the work of Molander and co-workers, we envisaged copper-catalysed β -boration of the newly formed α,β -unsaturated carbonyl with bis(pinacolato)diboron to produce the model boronic ester.³¹ While this procedure makes use of CyJohnPhos as phosphine ligand and sodium *tert*-butoxide as base, we found that alteration of these reagents to available xantphos and potassium *tert*-butoxide unproblematic. Following a small sample screening of catalyst and ligand loading amounts, optimal conditions were found and implemented in a Schlenk vessel to afford compound **44** as a clear oil in a yield of 81%. To our knowledge, this is the first instance of this reaction being performed on an aromatic unsubstituted acrylamide group using a copper catalyst.

We were intrigued by the publication of a two-step procedure for the deprotection of alkylpinacolyl boronate esters developed by Santos et al.³² This method, which involved transesterification using diethanolamine followed by hydrolysis with dilute hydrochloric acid, was attractive to us due to its varied functional group tolerance and ease of product purification (simple filtration). This allowed us to avoid the other relatively harsh routes found in literature.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



Scheme 7.8: Synthesis of model boron containing compounds **44** – **46**. Reagents and conditions: i) acryloyl chloride (1.1 equiv.), DIPEA (1.2 equiv.), DCM, $-40\text{ }^{\circ}\text{C}$ – $-20\text{ }^{\circ}\text{C}$, 2 h; ii) bis(pinacolato)diboron (1.2 equiv.), CuCl (10 mol%), potassium *tert*-butoxide (30 mol%), xantphos (10 mol%), EtOH, rt, 6 h; iii) diethanolamine (1.1 equiv.), diethyl ether, rt, 8 h; iv) 0.1 M HCl, diethyl ether, rt, 2 h.

Solution of boronic ester **44** in diethyl ether, followed by treatment with a slight excess of diethanolamine, facilitated precipitation of the diethanolamine salt adduct within the ethereal solution over an 8-hour period. As prescribed, filtration and successive washing of the white precipitate with diethyl ether afforded **45** in an acceptable yield. Lastly, subsequent hydrolysis was affected by suspension of **45** within a diethyl ether solution and the addition of 0.1 M HCl. Dilution of the reaction mixture with water, followed by extraction with ethyl acetate, afforded the boronic acid **46** with no need for further purification, thereby highlighting the simplicity and effectiveness of this methodology.

A comparison of the ^{11}B NMR spectra of each boron containing model compound is illustrated on the following page in **Figure 7.11**. When running boron NMR spectroscopy experiments with regular NMR tubes, broad interference signals are typically seen in any region upfield from approximately δ 20.0 ppm, as these tubes are made from borosilicate glass and consequently contain boron. It is therefore preferential to make use of quartz tubes, which are more expensive and fragile. All the collected ^{11}B NMR spectra were self-run in quartz tubes with chemical shifts referenced to 15% $\text{BF}_3 \cdot \text{OEt}_2$ in CDCl_3 , which served as the internal standard.

The NMR spectra of alkylpinacolyl boronate ester **44** exhibits a broad boron signal (red) appearing at δ 34.5 ppm, typical of this functional group with signals exemplified to occur between approximately δ 32.0 and 35.0 ppm.^{33, 34} Following transesterification to the diethanolamine salt adduct **45**, we see a large upfield shift of δ 21.5 ppm in the boron signal (blue) which appears as a broad peak at δ 13.0 ppm. This shift can be attributed to the cluster of electron density surrounding the now negatively charged boron atom with four bonds, creating a substantial shielding effect and the movement of the signal to an upfield position. Analysis of the ^{11}B NMR spectroscopic values of the diethanolamine adducts reported in the literature procedure reveals a chemical shift range between δ 10.0 and 13.0 ppm, corroborating the successful isolation of intermediate **45**.³² Lastly, mild hydrolysis of the adduct with 0.1 M HCl to provide boronic acid **46** reverts the upfield shift witnessed in the previous spectra, with the respective boron signal (green) arising at δ 33.1 ppm. Once again, appearance at this chemical shift is supported by the study undertaken in the literature, with typical boronic acid shifts lying anywhere between δ 19.0 and 35.0 ppm.³²

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

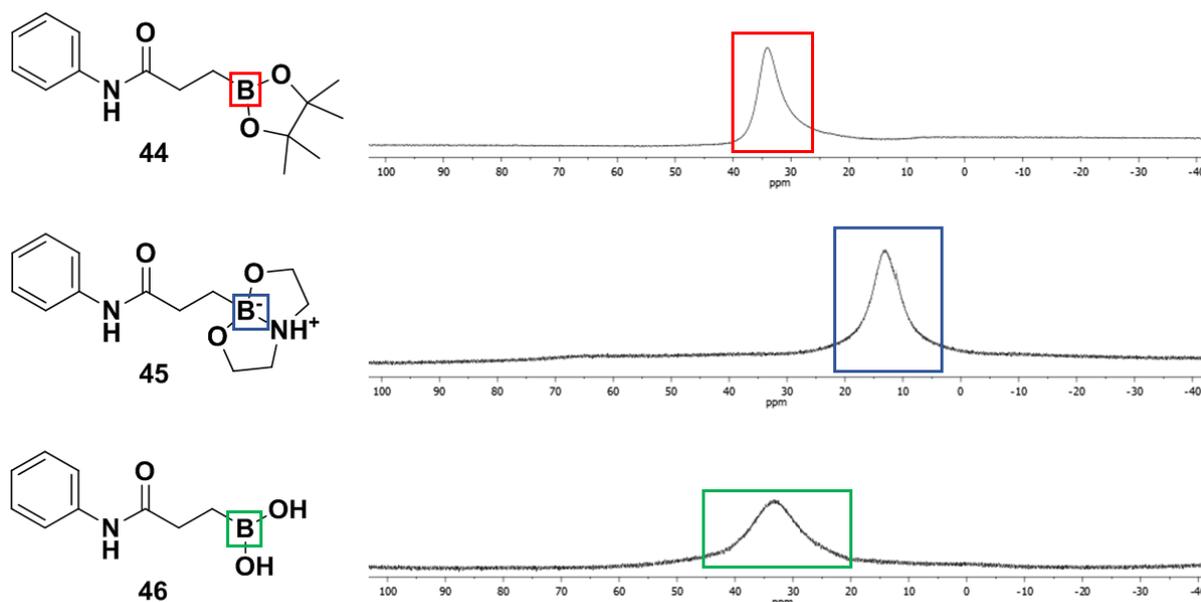
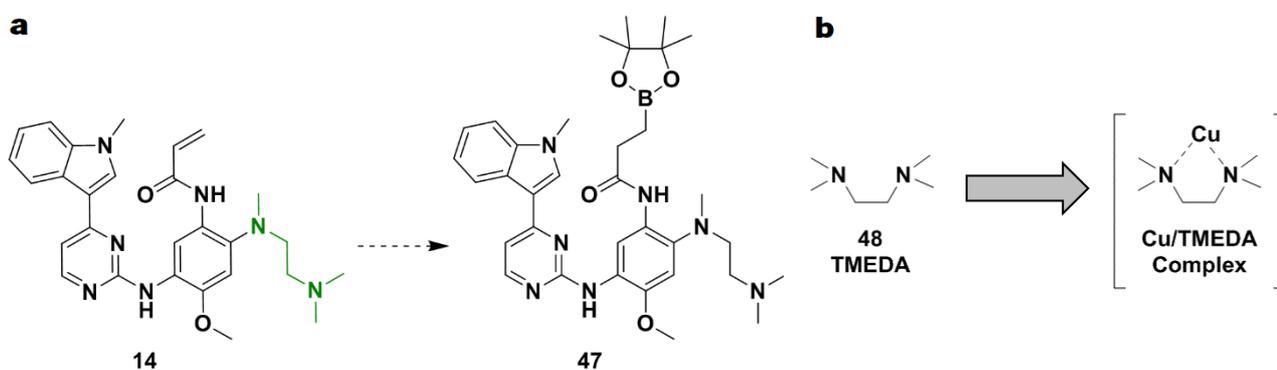


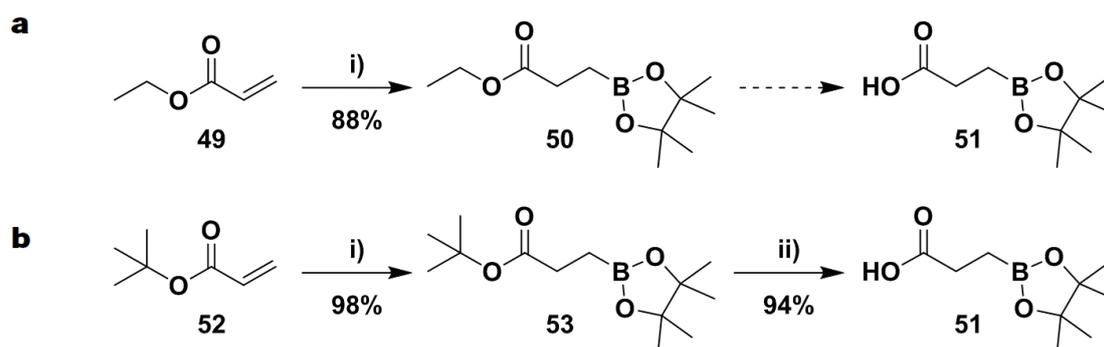
Figure 7.11: Comparison of the ^{11}B NMR spectra of model compounds **44** – **46**.

With the model synthesis of boronic ester and acid compounds completed, we focused our attention on the application of the derived reaction parameters towards final compound synthesis. The first of these steps entailed copper-catalysed β -boration of the previously synthesised osimertinib (**14**) with bis(pinacolato)diboron, illustrated in **Scheme 7.9**. To our surprise, employment of the model reaction conditions resulted in no reaction taking place and a full return of starting material **14**. We hypothesised that the root of this inactivity lay within the use of a copper catalyst and the coordinative nature of the osimertinib solubilising group (green, **14**), which is structurally related to the ligand tetramethylethylenediamine **48** (TMEDA) shown in **Scheme 7.9b**. TMEDA is widely employed as a ligand for metal ions and is able to form stable complexes with many metal halides such as copper(I) chloride.³⁵ In such complexes, TMEDA serves as a bidentate ligand (**Scheme 7.9b**). As such, we believe that the CuCl catalyst utilised in our β -boration reaction was potentially coordinating/complexing to the osimertinib solubilising group and therefore unable to participate in the reaction. Frustratingly, further attempts to overcome complexation using stoichiometric amounts of copper catalyst, in tandem with elevated temperatures and different solvents, met with the same result. We chose to abandon installation of the pinacol boronic ester group in this manner and proposed a new synthetic route via an acid coupling reaction, analogous to that used for the phosphonate esters **39** and **41**.



Scheme 7.9: a) Attempted synthesis of pinacolyl boronate ester **47** and b) TMEDA (**48**) and its copper complex.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



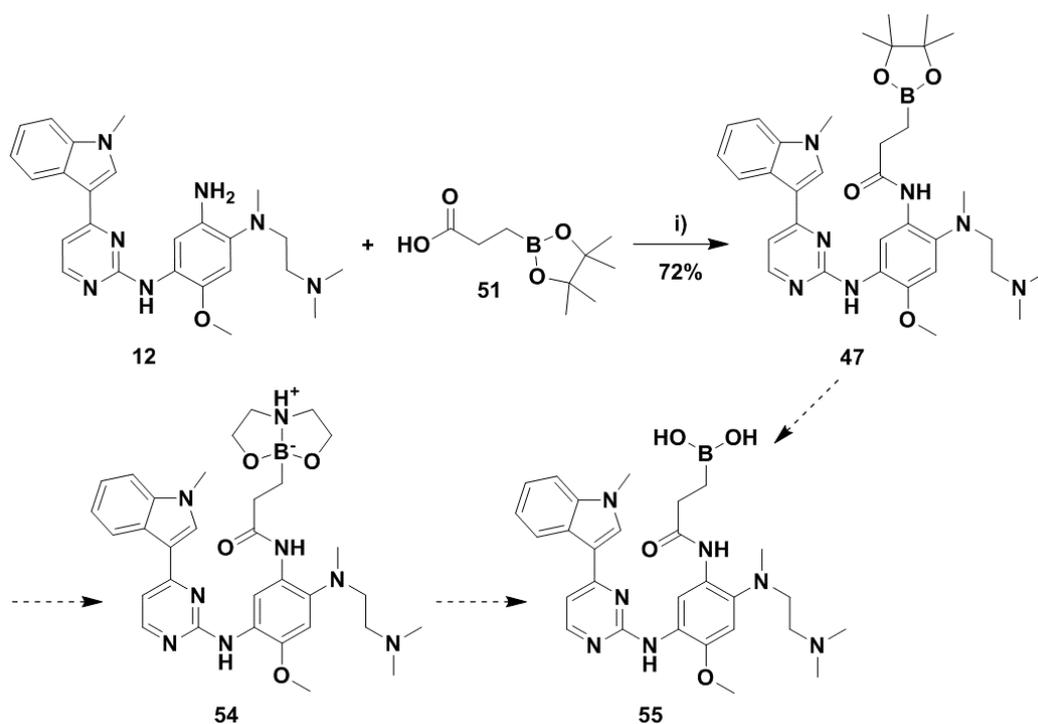
Scheme 7.10: a) Attempted synthesis and b) synthesis of carboxylic acid pinacolyl boronate ester fragment **51**. Reagents and conditions: i) bis(pinacolato)diboron (1.2 equiv.), CuCl (10 mol%), potassium *tert*-butoxide (30 mol%), xantphos (10 mol%), EtOH, rt, 6 h; ii) bis(pinacolato)diboron (1.05 equiv.), CuCl (5 mol%), potassium *tert*-butoxide (10 mol%), xantphos (5 mol%), THF/MeOH (20:1), rt, 5 h; iii) DCM/TFA (3:1), rt, 6 h.

Execution of our alternative synthetic strategy required the successful synthesis of the carboxylic acid fragment **51** (**Scheme 7.10**), which could then undergo amide bond formation with driving scaffold **12**. Accordingly, ethyl acrylate (**49**) was subjected to copper-catalysed β -boration to afford the pinacolyl boronate ester **50** in excellent yield, as illustrated in **Scheme 7.10a**. Disappointingly, the attempted selective hydrolysis of ethyl ester **50** proved impractical, as we were unable to extract and isolate the formed carboxylic acid **51** from the aqueous phase during work-up neutralisation. This included the use of various metal salts (LiBr, LiOH, KOH, NaOH) at adjusting temperatures, followed by treatment with acid (HCl, H₂SO₄) - all to no avail.

To circumvent the obstacle posed by post-hydrolysis neutralisation and extraction from the aqueous phase, we envisaged replacement of the ethyl ester with that of a *tert*-butyl ester. Facile and effective cleavage of a *tert*-butyl ester functional group by stirring in trifluoroacetic acid is well documented and we believed we could implement this to our advantage.³⁶ While sceptical at the prospect of the pinacol boronic ester surviving this treatment, we maintained the venture was worth a try. Interestingly, the discovery of a similar literature procedure for β -boration of the commercially *tert*-butyl acrylate (**52**) resulted in a near quantitative yield of **53**, as illustrated in **Scheme 7.10b**.³⁷ Solvent exchange from ethanol to a tetrahydrofuran/methanol mixture, smaller catalytic amounts of reagents and the order of addition were the major contributing factors to this yield increase. To our delight, treatment of the newly furnished **53** with trifluoroacetic acid whilst stirring in DCM, followed by simple removal of the solvent *in vacuo*, afforded the desired fragment **51** in exceptional yield with no further purification required. Survival of the pinacol group and the successful synthesis of carboxylic acid **51** was confirmed by HRMS, ¹H, ¹³C and ¹¹B NMR spectroscopy, with the latter revealing the appearance of a boron signal at δ 34.5 ppm.

Illustrated in **Scheme 7.11** on the next page, efforts towards the synthesis, isolation and purification of **47** proved to be an incredibly frustrating endeavour. The use of an array of conventional amide-coupling reagents including CDI, various carbodiimides (DCC, DIC, EDC) and PyBOP was perceived to affect successful amide bond formation as visualised by TLC. However, following the work-up procedure, subsequent purification by flash column chromatography would consistently result in the loss of this product spot, presumably due to degradation on the silica column.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



Scheme 7.11: Synthesis of pinacolyl boronic ester **53** and attempted hydrolysis towards boronic acid **55**. Reagents and conditions: i) **51** (2 equiv.), HOBT·H₂O (2 equiv.), EDC·HCl (2 equiv.) DCM, 0 °C – rt, 8 h.

As column chromatography was not tolerated, we devised a synthetic procedure that could afford the pure product using only the work-up as the method of purification. Following numerous attempts, the optimal parameters of two equivalents of the carboxylic acid **51**, EDC hydrochloride and HOBT hydrate in DCM were established. Use of an excess of reagents was required to ensure complete conversion to the desired product. During the work-up procedure, washing of the organic phase with water ensured removal of the water-soluble EDC coupling agent and the activating additive HOBT. This was followed by successive washing with a 1 M NaOH solution to remove any excess carboxylic acid **51**. Gratifyingly, removal of the solvent *in vacuo* afforded the coupled product **47** as a white foam in a yield of 72%, allowing us to proceed towards hydrolysis of the pinacol boronic ester.

Unfortunately, utilisation of diethanolamine in diethyl ether to facilitate hydrolysis of boronic ester **47** was met with no reaction and a return of the starting material. Integral to precipitation of the intermediate adduct **54**, the requirement of diethyl ether as solvent was also found to be the main impediment, due to the insolubility of **47**. While replacement of diethyl ether with other non-polar solvents such as tetrahydrofuran, toluene and chloroform did permit the solution of **53**, no precipitation of any product was found to occur, highlighting the necessity of an ethereal solution. Furthermore, any attempts at purification of the formed polar products using column chromatography resulted in the invariable degradation and loss of compound experienced in the synthesis of **47**.

In a final effort, direct hydrolysis of the boronic ester **53** was attempted using the conventional harsh method of heating in concentrated hydrochloric acid. To our dismay, this resulted in the formation of numerous by-products, with purification by column chromatography plagued by product degradation. Due to time constraints, it was decided at this point to submit **47** for biochemical and cellular evaluation to ascertain whether there was any merit in pursuing the synthesis of boronic acid **55**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

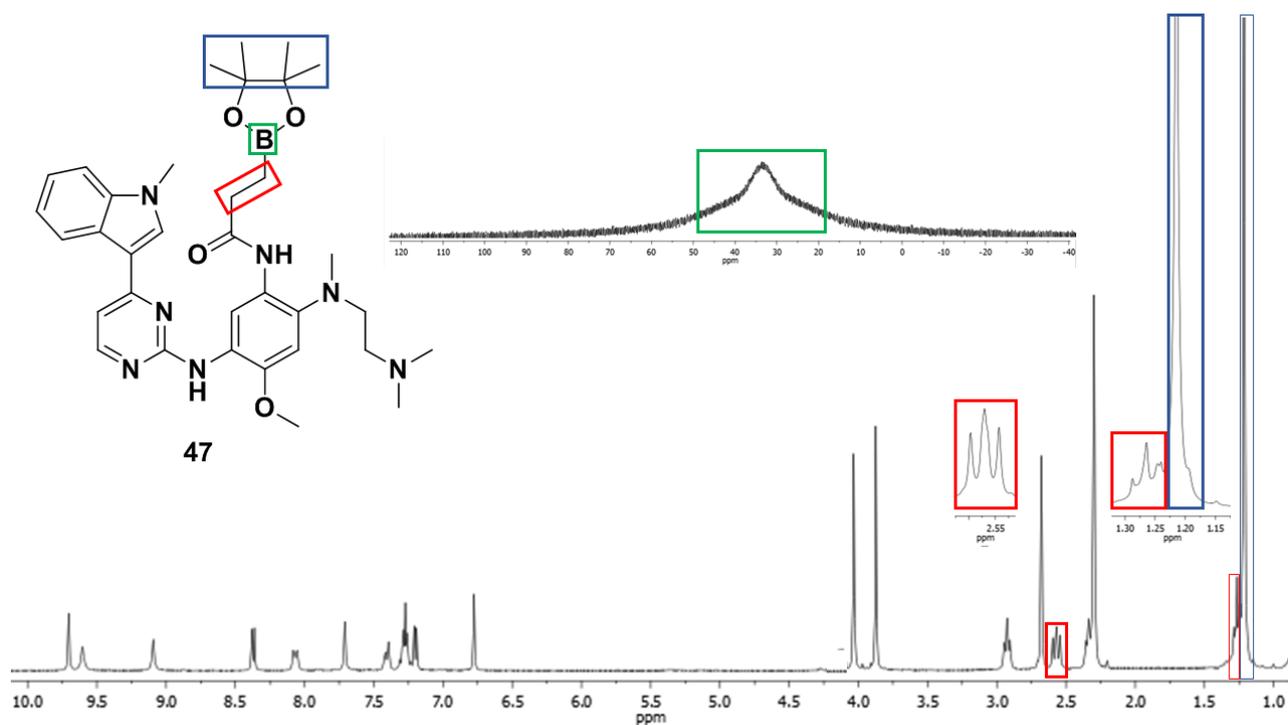
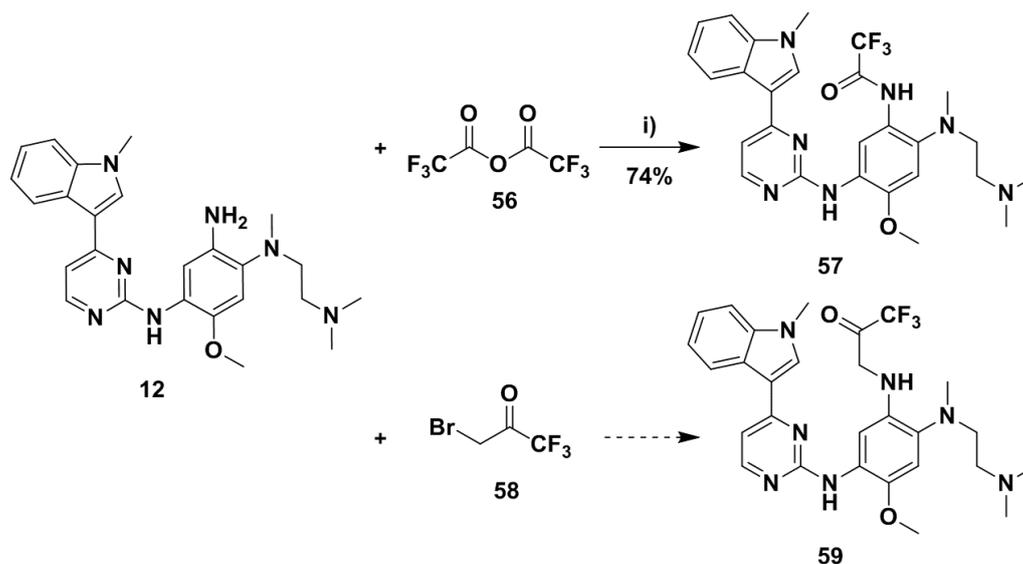


Figure 7.12: ^1H and ^{11}B NMR spectra of final compound **53**.

The ^1H and ^{11}B (inset) NMR spectra of **47** may be found in **Figure 7.12** above. As full characterisation of the driving group scaffold has already been undertaken in the ^1H NMR spectroscopic analysis of compound **17** (**Figure 7.8**), we will focus solely on the newly incorporated elements of the boronic ester. Examination of the most upfield region reveals the presence of a large singlet at δ 1.21 ppm integrating for 12H, ascribed to the four methyl groups of the pinacol moiety (blue). This singlet was flanked by a nearby triplet at δ 1.26 ppm which integrated for 2H. Working from the NMR spectra of precursory structure **12**, we assigned this triplet to one of the ethyl bridge (red) methylene protons. Assignment of the other unresolved methylene linker was expedited through comparison of the relevant J-coupling values. Appropriately, the triplet with a J-coupling value of 6.2 Hz appearing at δ 2.57 ppm was prescribed to the remaining methylene (red), as this value was identical to that of the previously assigned triplet. Additionally, the neighbouring triplets J-coupling values were both found to be 5.9 Hz, validating this assignment. Lastly, the ^{11}B NMR spectra revealed the presence of a broad singlet at δ 33.3 ppm, corresponding to the boron atom signal of the pinacol ester (green).

Assessment of the unassigned signals within the ^1H NMR spectrum presented the correct integration amount for compound **47**, with all protons accounted for and peak splitting patterns to be in good order. While not displayed for interpretation, all carbons were correctly accounted for in the corresponding ^{13}C NMR spectra, with the pinacol group and newly installed amide carbonyl carbon particularly evident. Use of HRMS analysis presented a $[\text{M}+\text{H}]^+$ ion of 628.3775, with a calculated mass of 628.3783, providing final confirmation for the successful synthesis of compound **47**. While this brought an end to the synthesis of our boron-containing final compounds, further considerations towards the synthesis of this class of inhibitors will be discussed in the future work section of this chapter (**Section 7.7**).

7.4.3.5 Trifluoromethyl ketones



Scheme 7.12: Synthesis of trifluoromethyl ketone **57** and attempted synthesis of trifluoromethyl ketone **59**. Reagents and conditions: i) TFAA (2 equiv.), DIPEA (1.2 equiv.), THF, $-10\text{ }^{\circ}\text{C}$ – rt, 4 h.

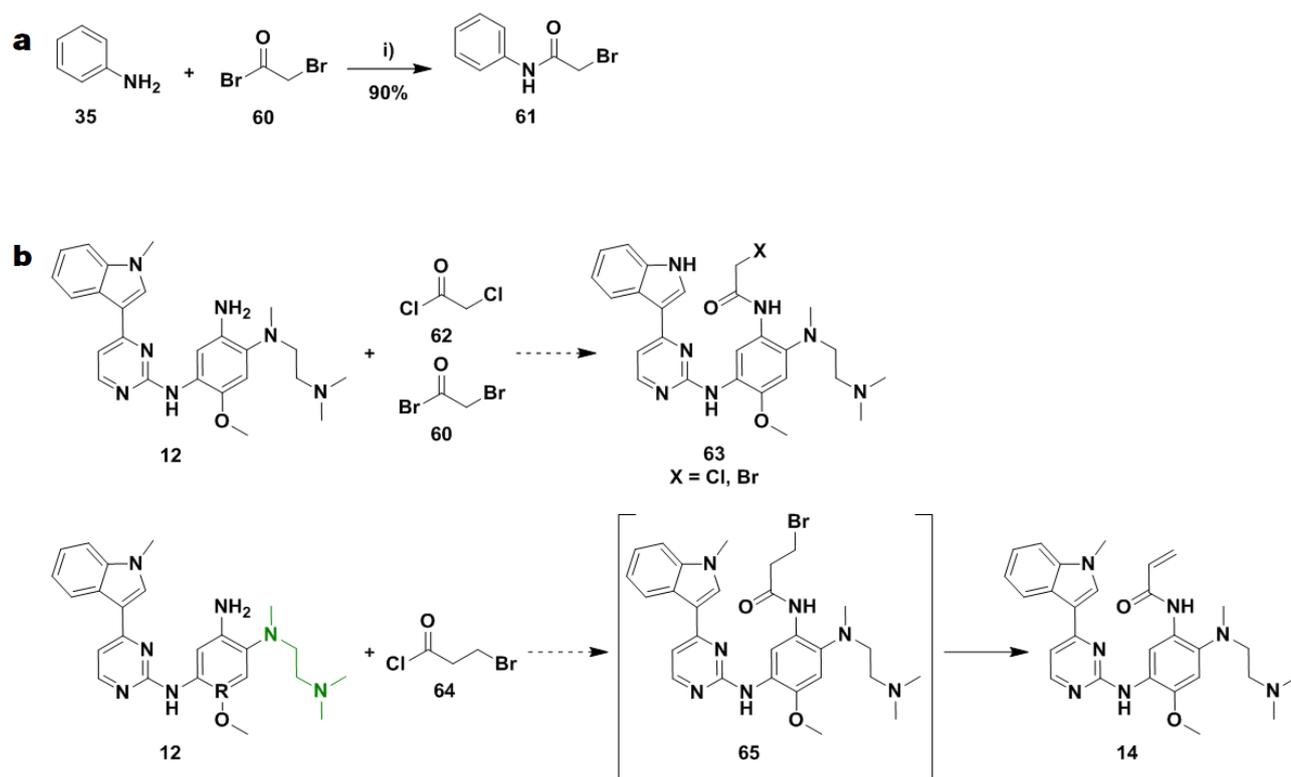
The synthesis of our trifluoromethyl ketone containing final compounds, illustrated above in **Scheme 7.12**, was unambiguous and will therefore be discussed briefly. Treatment of the osimertinib-derived scaffold **12** with trifluoroacetic anhydride at low temperature, with DIPEA serving as base, facilitated production of trifluoromethyl ketone compound **57** in a yield of 74%. This compound was fully characterised using IR, HRMS and NMR spectroscopic techniques, verifying its successful synthesis. Acquisition of the commercially available electrophilic fragment **58** and its subsequent reaction with **12** was met with the formation of a multitude of inseparable by-products, even at reaction temperatures of $-78\text{ }^{\circ}\text{C}$. We believe the generation of by-products stems from the inherent reactivity of this fragment (two electrophilic sites) and the availability of nucleophilic amines within scaffold **12** to promote undesired secondary reactions. Consequently, we were unable to successfully isolate trifluoromethyl ketone **59** and continued towards the synthesis of our remaining serine-targeting inhibitors.

7.4.3.6 α -, β -Haloketones

Due to the difficulties experienced in the previous chapters synthesis of α -, β -haloketones, we wished to investigate incorporation of this electrophile on a model system, shown on the following page in **Scheme 7.13a**. Treatment of **35** with the previously distilled bromoacetyl bromide (**60**), at low temperatures generated by an acetonitrile/dry ice bath, furnished the model α -haloketone **61** in 90% yield. This was unanticipated, as previous encounters with this class of compounds had displayed their highly reactive and unstable nature.

Discouragingly, applications of these reaction conditions to scaffold **12** with various electrophiles (**60**, **62**) to afford the α -haloketones (**63**) were met with failure, as illustrated in **Scheme 7.13b**. We found the synthesis of the α -haloketones **63** to mirror the circumstances of the previous chapter. Whilst maintaining the low reaction temperatures, it appeared as though single product formation had taken place when visualised by TLC analysis. However, upon warming to room temperature, the previously identified product spot degraded into a profusion of by-products, making the attempted the isolation and purification of these compounds impossible.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



Scheme 7.13: a) Synthesis of model α -haloketone **61** and b) attempted synthesis of α -, β -haloketones **63** and **65**. Reagents and conditions: i) **60** (1.1 equiv.), DIPEA (1.2 equiv.), DCM, $-40\text{ }^{\circ}\text{C} - -20\text{ }^{\circ}\text{C}$, 2 h.

In an attempt to diminish the reactivity of the system, the temperature was lowered to $-78\text{ }^{\circ}\text{C}$, in conjunction with removal of any external base. Unfortunately, this led to same outcome of degraded and inseparable crude reaction mixtures and the inability to furnish **63**.

Focussing on the synthesis of β -haloketone **65**, acylation of **12** with electrophile **64** also led to the formation of a single spot as per visualisation by TLC analysis. In contrast to the α -haloketone system, warming of the crude reaction mixture to room temperature and subsequent work-up procedure did not lead to compound degradation, allowing us to isolate and purify the crude product. To our surprise, spectroscopic analysis by HRMS and NMR spectroscopy experiments yielded spectra matching that of osimertinib (**14**), postulated to have occurred through the unexpected elimination of the desired **65**. A search through the literature revealed a recently developed procedure employing a similar two-step acylation/elimination sequence.³⁸ This was accomplished using the chloro-variant of **64**, followed by treatment with Et_3N in acetonitrile heated under reflux.

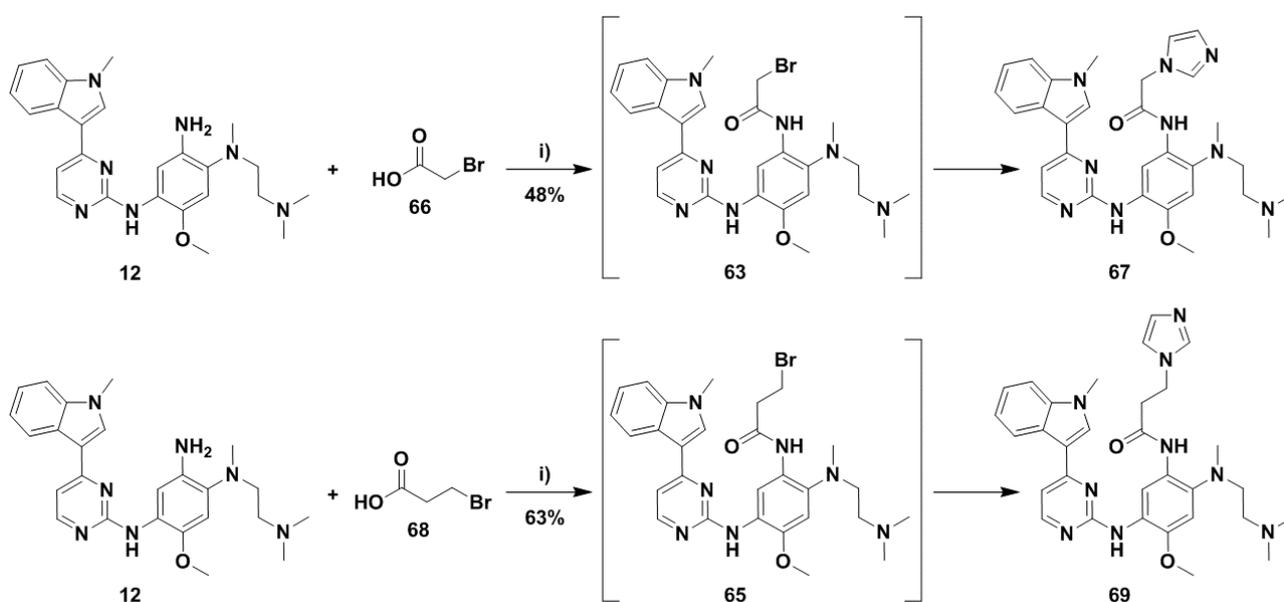
Intriguingly, removal of all base to circumvent spontaneous elimination in the acylation of **12** with **64** generated a single polar spot when visualised by TLC. Curiously, upon subjection to aqueous work-up and neutralisation, this spot shifted to the R_f value corresponding to that of osimertinib (**14**), which it was later confirmed as by NMR spectroscopic analysis. Witnessing these transformations, we hypothesised that the solubilising group (green) was serving as an intra-molecular base in these reactions. This is supported by the basic nature of the disubstituted amine groups and the rotational freedom of the solubilising group, permitting a proximal location to the β -haloketone. Therefore, in the absence of any external base, we believe that the trimethylethylenediamine solubilising group (green) facilitates elimination of the formed intermediate **65**, producing the hydrobromide salt of osimertinib. This would correlate to the polar spot visualised during TLC analysis and its conversion into **14** following neutralisation.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

While speculative, this would provide an explanation for the problems encountered in the synthesis of the α - and β -haloketone-containing final compounds in both this chapter and the previous one. This hypothesis is further corroborated in the attempted synthesis of α - and β -haloketones through an amide coupling reaction, resulting in the unforeseen formation of imidazole bearing final compounds. The discussion of this reaction sequence and its relation to our hypothesis will continue in the following section.

7.4.3.7 Imidazole Compounds

During the synthetic exploration of the α - and β -haloketones **63** and **65**, we attempted amide coupling reactions between driving scaffold **12** and the carboxylic acid electrophiles **66** and **68**, shown below in **Scheme 7.14**. For the initial reaction between **12** and **66**, we utilised the previously established conditions of excess carboxylic acid and CDI as coupling reagent. During this reaction, we observed the preliminary formation of a single product spot through use of TLC analysis. Interestingly, this was followed by the slow conversion of this spot into another more polar product over a prolonged period. Accordingly, reaction times were extended to allow for full conversion to this unidentified adduct. Fascinated by the unconventional stability following isolation after work-up, purification of the unknown compound was achieved using flash column chromatography. Subsequent IR, HRMS and ^1H and ^{13}C NMR experiments pointed towards the unanticipated, but welcome synthesis of imidazole-containing compound **67** in an unoptimized yield of 48%. Following the initial amide coupling to produce intermediate α -haloketone **63**, it was evident that nucleophilic attack by excess imidazole (released during this amide bond formation) at the electrophilic α -position of **63** resulted in formation of the stable compound **67**. Exceptionally, application of the same coupling parameters between **12** and **68** led to the generation of compound **69** in an improved yield, despite the less reactive β -halide electrophile centre. This compound was also fully characterised using IR, HRMS and NMR spectroscopy.



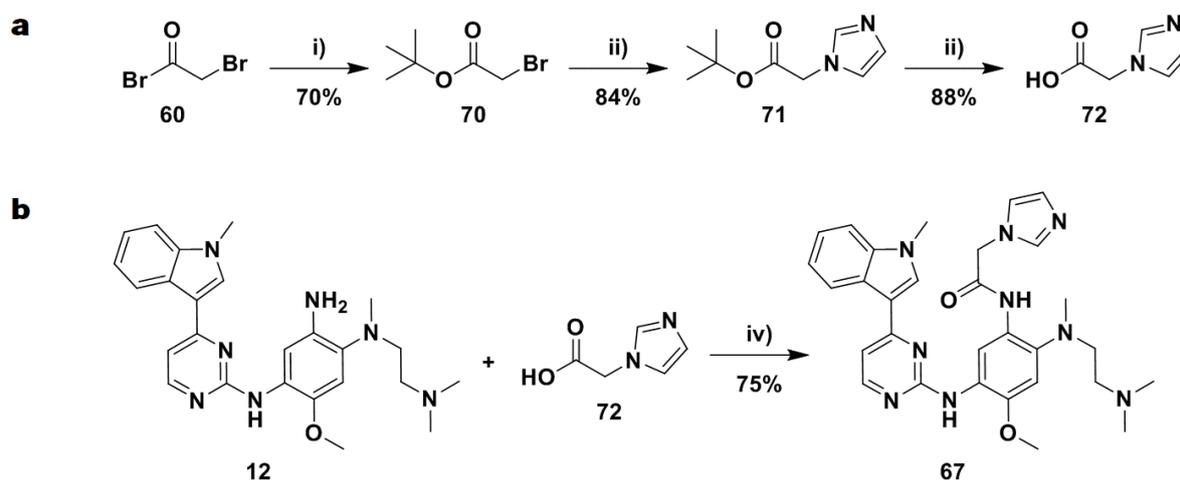
Scheme 7.14: a) Synthesis of model α -haloketone **61** and b) attempted synthesis of α -, β -haloketones **63** and **65**. Reagents and conditions: i) **66** or **69** (1.85 equiv.), CDI (1.85 equiv.), DCM, 0 °C – rt, 12 h.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

As no external source of base was present in either reactions, we are presented with two potential candidates which may have fulfilled this role and expedited the substitution reaction. As we utilised an excess of CDI in the coupling reaction, the remaining imidazole that did not participate in the substitution reaction could have functioned as a base. Alternatively, building upon the speculations and reasons outlined in the previous section, the trimethylethylenediamine solubilising group could have acted as an intra-molecular base. Appraisal of the relevant pK_a values of the conjugate acids of imidazole and TMEDA, which is structurally related to our solubilising group, shows values of 6.95 and 8.97 respectively.^{35, 39} Based on these numbers, the solubilising group is the more basic, aiding weight to our rational of its potential role as intra-molecular base.

While no studies were undertaken to investigate these claims, acid coupling reactions could be performed on a model system, such as aniline, which excludes the solubilising group. Use of excess CDI, in the presence or absence of external base, will provide pertinent information needed to ascertain the roles of the reactants. Finally, we could find no occurrence of this reaction in our perusal of the literature, highlighting it as a new method for potential one-pot installation of these imidazole-bearing functional groups.

Sceptical of the above discussed transformations, we elected to proceed with what would have been our original synthetic route towards our imidazole-containing final compounds. The route, displayed below in **Scheme 7.15**, emulated the synthesis developed for the boron-containing carboxylic acid **50** shown in **Scheme 7.10b**. In doing so, we would take advantage of the *tert*-butyl ester functional group and its facile hydrolysis using trifluoroacetic acid. Moreover, successful isolation of reactive fragment **72** would cater for convergent synthesis in future synthetic libraries.



Scheme 7.15: a) Synthesis of imidazole carboxylic acid **72** and b) imidazole-containing final compound **67**. Reagents and conditions: i) **60** (1.1 equiv.), *tert*-butanol (1 equiv.), pyridine (1.5 equiv.), DCM, $-10\text{ }^\circ\text{C}$ – rt, 4 h; ii) **70** (1.1 equiv.), imidazole (1 equiv.), Cs_2CO_3 (1.2 equiv.), DMF, $0\text{ }^\circ\text{C}$ – rt, 4 h; iii) DCM/TFA (3:1), rt, 6 h; iv) **72** (2 equiv.), HOBt·H₂O (2 equiv.), EDC·HCl (2 equiv.) DCM, $0\text{ }^\circ\text{C}$ – rt, 12 h.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Initial attempts to obtain compound **70** (**Scheme 7.15a**) using Steglich esterification were met with low yields and tedious reaction times. We therefore revised our synthetic strategy, utilising an acylation reaction between the highly reactive bromoacetyl bromide **60** and *tert*-butanol, with pyridine serving as base.⁴⁰ Following a simple work-up procedure, which included a dilute acid wash and extraction with diethyl ether, compound **70** was afforded in good yield with no further purification required. The subsequent substitution reaction of **70** with imidazole proceeded smoothly at room temperature using caesium carbonate as base in DMF.⁴¹ This furnished the penultimate product **71** as a tan solid in a yield of 84%. Lastly, stirring of compound **71** with trifluoroacetic acid in DCM for 6 hours was followed by removal of the solvent *in vacuo* to produce the desired imidazole carboxylic acid **72** in excellent yield. This highlights the robust and effective nature of this methodology in cleaving *tert*-butyl ester functional groups.

With fragment **72** in hand, we set out towards the ensuing amide formation with driving group scaffold **12**, shown in **Scheme 7.15b**. Employing the optimized conditions developed for our boron-containing final compounds, two equivalents of **72**, EDC hydrochloride and HOBt hydrate was stirred with **12** at ambient temperature for 12 hours. Following facile consumption of the starting material, the crude product was isolated and purified by flash column chromatography. To our delight, scrutiny of the compound by HRMS and NMR spectroscopy revealed identical spectra to that obtained in **Scheme 7.14** utilising CDI, confirming the successful synthesis of **67** in an improved yield of 75%. To discern this, the ¹H NMR spectra of both the *tert*-butyl ester imidazole building block **71** and final compound **67** will be examined and interpreted. Fragment **71** was chosen over the carboxylic acid **72**, as the latter was only soluble in deuterated methanol.

Inspection of the basic ¹H NMR spectra of **71**, illustrated on the following page in **Figure 7.13a**, provides structural confirmation for the successful synthesis of this intermediate. While not relevant in comparison with **67**, we see a large singlet furthest upfield at δ 1.45 ppm integrating for 9H which was easily ascribed to the *tert*-butyl group (purple) of **71**. Moving downfield, a lone peak corresponding to the methylene linker (red) can be found at δ 4.56 ppm. This left three signals in the aromatic region which all appeared as relatively broad singlets and integrated for 1H each. Based on the deshielding effects of the neighbouring nitrogen atoms, the 2-position proton (blue) of the imidazole was assigned to the peak farthest downfield at δ 7.47 ppm. By process of elimination, the final two signals at δ 6.93 and 7.07 ppm were correlated to the remaining imidazole protons (green). Surprisingly, these signals did not appear as the anticipated doublets but rather as broad singlets. This peak broadening can possibly be explained by the tautomeric and conformational properties of the imidazole group.

Comparative analysis with the ¹H NMR spectra of **67** (**Scheme 7.13b**) reveals the emergence of four new peaks in positions correlating with those found in the spectra of **71** thereby supporting successful installation of the imidazole carbonyl functional group. As with the **71**, we see a solitary singlet that has surfaced at δ 4.87 ppm which integrates for 2H, clearly indicative of the methylene group (red) which serves as a linker between the carbonyl and imidazole moiety. Scrutiny of the aromatic region reveals a similar peak pattern to that found in the ¹H NMR spectra of **71**, namely three broad singlets that integrate for 1H each. Additionally, these peaks were not evident in the spectra of the precursory driving group structure **12**. By these factors and logical deduction, the two peaks at δ 7.07 and 7.17 ppm were assigned to the neighbouring imidazole protons (green), while the singlet at δ 7.63 ppm was ascribed to the final 2-position proton (blue) of the imidazole functionality.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

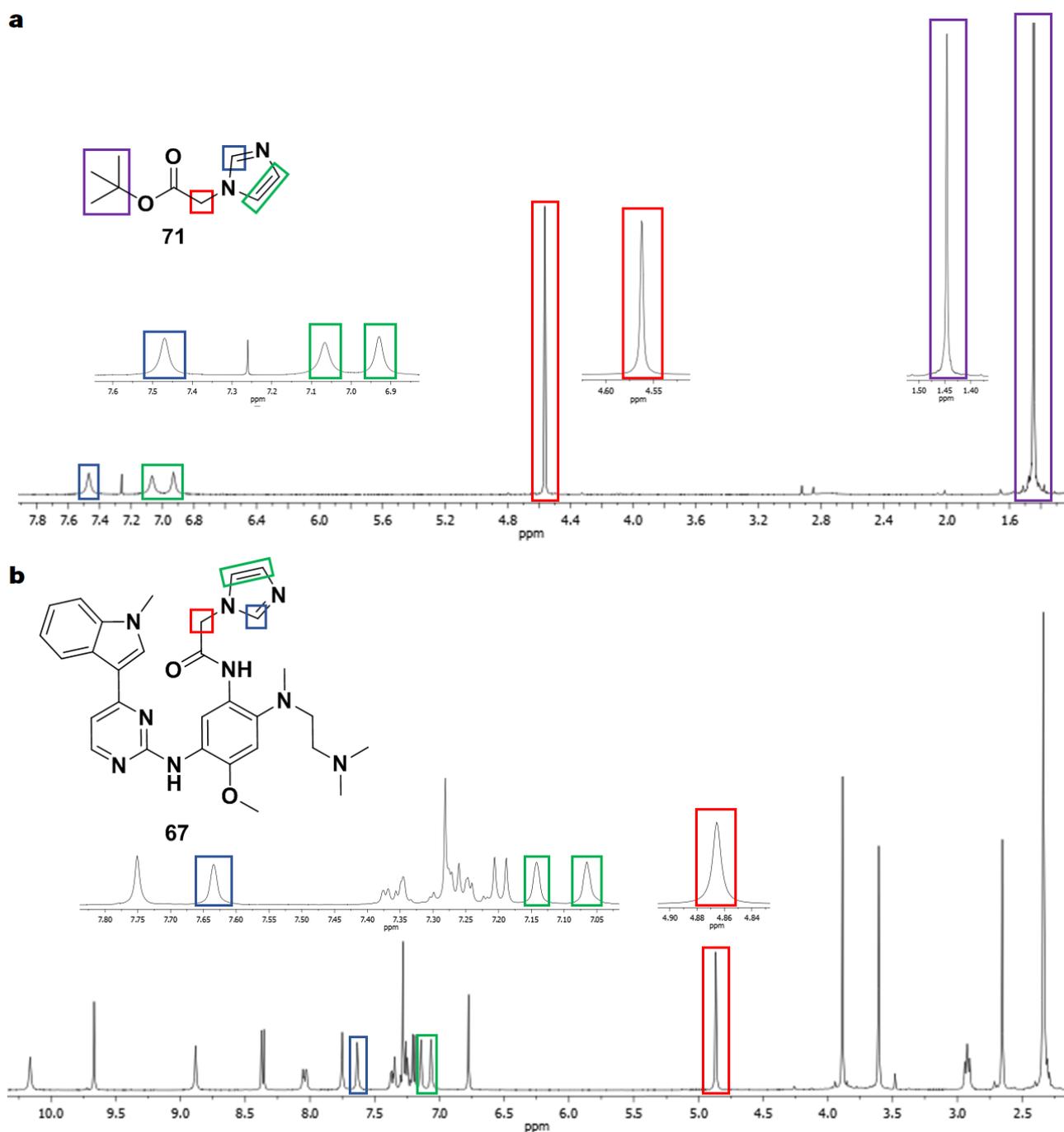
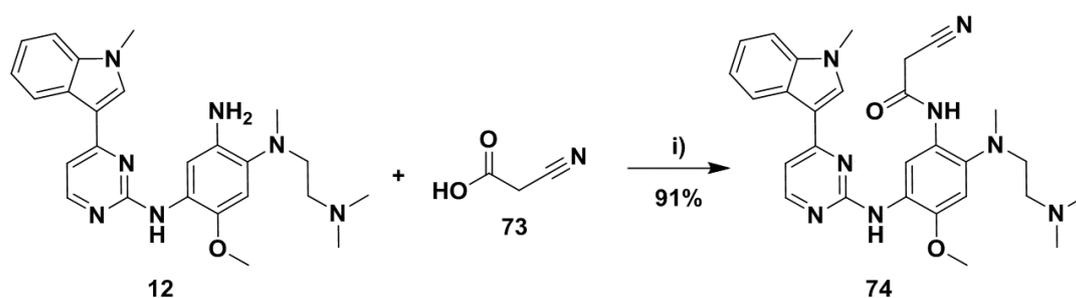


Figure 7.13: ^1H NMR spectra of a) imidazole building block **71** and b) final compound **67**.

Assessment of all other signals within the ^1H NMR spectra presented the correct integration amount for compound **67**, with all protons accounted for and peak splitting patterns to be in good order. The same held for the ^{13}C NMR spectrum. Use of HRMS analysis presented a $[\text{M}+\text{H}]^+$ ion of 554.3007, with a calculated mass of 554.2992, providing final confirmation for the successful synthesis of compound **67**.

While not described, all signals within the ^1H and ^{13}C NMR spectra of the structurally related final compound **69** were correctly accounted for. Furthermore, compound **69** displayed a $[\text{M}+\text{H}]^+$ ion of 568.3158, with a calculated mass of 568.3148, verifying its effective isolation and concluding the synthesis for our imidazole-based final compounds.

7.4.3.8 Nitriles



Scheme 7.16: Synthesis of nitrile-containing final compound **74**. Reagents and conditions: i) **73** (2 equiv.), HOBT·H₂O (2 equiv.), EDC·HCl (2 equiv.) DCM, 0 °C – rt, 5 h.

Only one compound bearing a nitrile functionality was synthesised in our serine-targeting library of inhibitors, the synthesis of which is displayed above in **Scheme 7.16**. Application of the optimized acid-coupling parameters between **12** and cyanoacetic acid (**73**) afforded final compound **74** effortlessly in excellent yield. Full characterisation by IR, HRMS and NMR spectroscopy - providing a [M+H]⁺ ion of 513.2727 with a calculated mass of 513.2726 - validated the successful synthesis of the cyanoacetamide-containing final compound **74**. This led us to the last synthetic endeavour of this chapter involving exploration of dual functional warheads.

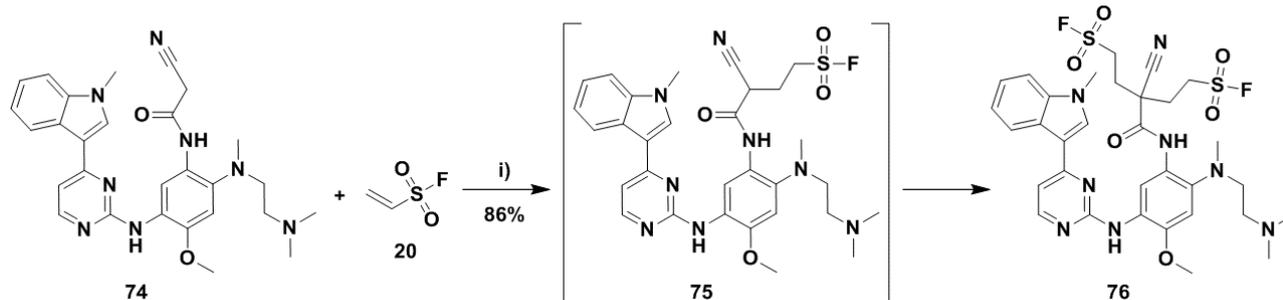
7.4.3.9 Dual Warheads

As outlined in **Section 7.3**, we conceived of several electrophiles that contain two functional groups, thereby able to simultaneously reversibly interact with and potentially covalently modify the mutated Ser797 residue. This design was based on the catalytic triad of serine proteases and our envisioned inhibitors incorporated the use of either a cyano- or imidazole-acetamide fragment in combination with a sulfonyl fluoride moiety. Accordingly, the synthesis relating to the cyano-based dual functional warhead is illustrated on the following page in **Scheme 7.17** and will be discussed first.

Our strategy towards the synthesis of these dual functional warheads was focused on the use of a Michael addition reaction. We proposed treatment of the previously synthesised **74** with mild base would facilitate deprotonation of the acidic cyanoacetamide α -proton, thereby creating a nucleophilic enolate at this position. Introduction of the appropriate electrophile, ESF (**20**), would then allow for the Michael addition reaction to take place and potentially afford dual warhead **75**. Due to its frequent use in deprotonation of 1,3-dicarbonyl compounds ($pK_a \sim 9 - 12$) and Michael addition reactions, we advocated potassium carbonate as a suitable base.⁴² As ESF contains two electrophilic sites (Michael acceptor and sulfonyl fluoride), we wished to avoid the competing reaction of nucleophilic substitution at the sulfonyl fluoride. However, Sharpless et al. demonstrated the remarkable stability and inertness of the generated sulfonyl fluoride, which only readily reacts with nucleophiles under specific conditions (aqueous solvent).⁴³ We therefore opted to employ acetonitrile as solvent, as it would facilitate solvation of the starting material **74** and provide low solubility/surface reaction with the commonly paired potassium carbonate.

Implementation of these parameters (**Scheme 7.17**), with an excess of **20** and potassium carbonate to ensure reaction completion, resulted in the observation of a single spot as per visualisation by TLC analysis.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



Scheme 7.17: Attempted synthesis of compound **75** resulting in isolation of final compound **76**. Reagents and conditions: i) **20** (1.85 equiv.), K₂CO₃ (2 equiv.), MeCN, -10 °C – rt, 8 h.

Enthused by this, we avoided a work-up procedure and the potential hazard of compound degradation and chose to directly purify the crude product using flash column chromatography. Demonstrated to be stable, the compound was subsequently subjected to initial analysis by ¹H and ¹³C NMR spectroscopy. Perplexingly, the processed spectra did not match that of the expected compound **75**, with proton integration and carbons signals equating to a larger amount than anticipated. The unknown compound was therefore submitted for HRMS analysis, presenting a [M+H]⁺ ion of 733.2407. This was strikingly similar to the calculated mass of 733.2402, which matched that of the di-substituted product **76**. At this point, we surmised that this could indeed be the structure of the isolated product. However, to authenticate this presumption, we obtained and processed the relevant 2D NMR spectra. Consequently, a variety of 2D NMR experiments were performed which included APT, COSY, HSQC and HMBC. The processed ¹H, ¹⁹F, ¹³C and portions of the HSQC and HMBC NMR spectra will be displayed and analysed.

The ¹H NMR spectrum of compound **76** may be found in **Figure 7.14a** on the following page. Visual inspection revealed the familiar structural signature of the osimertinib derived driving group, as well as the emergence of several unidentified signals. As full characterisation of the heterocyclic core had been previously undertaken, we focussed our efforts solely on the placement and identification of these unknown elements, which consisted of two sets of signals. Starting from the most upfield position, we saw the appearance of a triplet of doublets at δ 2.59 ppm which integrated for 2H. Scrutiny of the adjacent set of peaks at δ 2.34 ppm, which contained the dimethylamine (dark blue) element of the solubilising group, revealed an integration of 8H instead of 6H. Upon closer inspection, we see two sets of peaks flanking the outer edge of the large singlet. This signified an overlap with another triplet of doublets (green) which integrated for 2H. We proposed that these signals corresponded to the ethylene protons (green) of the doubly substituted sulfonyl fluoride groups which are situated closest to the quaternary carbon centre. Due to the unequivalence of these protons, we expected to observe two sets of signals and a peak splitting pattern similar to that displayed within the spectrum, thereby corroborating our assumption.

Moving downfield, we identified two multiplets at δ 3.34 and 3.53 ppm (red) which also integrated for 2H each. We ascribed this set of peaks to the neighbouring ethylene protons (red) of the double sulfonyl fluoride functionality. As with the adjacent ethylene (green), these protons are unequivalent and would therefore give rise to two signals. Furthermore, coupling to the fluorine atom of the sulfonyl fluoride would likely result in a more complex peak splitting pattern. Observation of both suppositions within the ¹H NMR spectrum validated the proposed structure of compound **76**. Lastly, the ¹⁹F NMR spectrum (inset) confirmed the presence of a sulfonyl fluoride (purple) moiety, with a triplet observed at δ 53.4 ppm – comparable to the spectra obtained for compounds **18** and **21** (**Figure 7.9**).

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

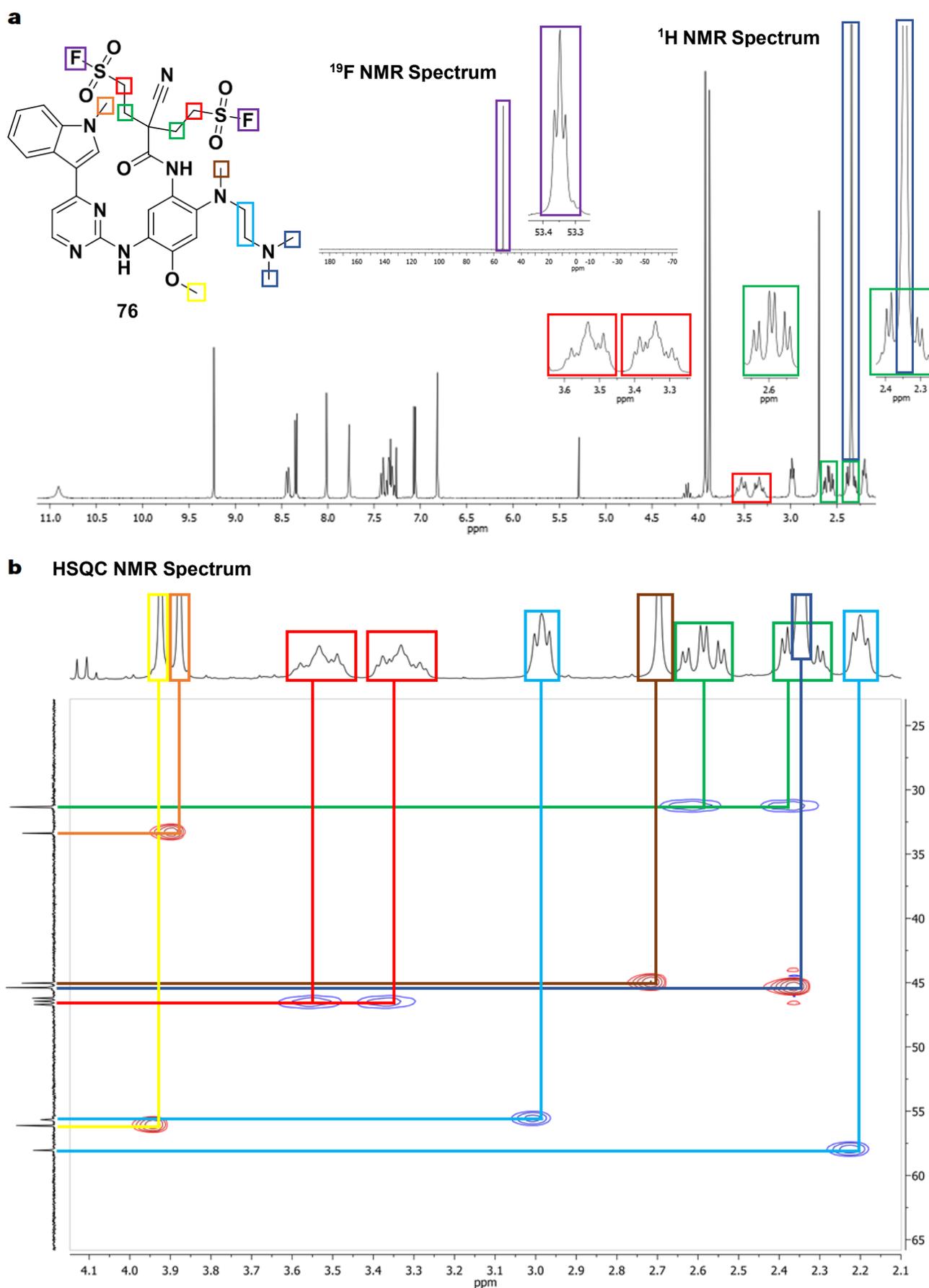


Figure 7.14: a) ^1H and ^{19}F (inset) NMR spectra and b) HSQC NMR spectra of final compound **76**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

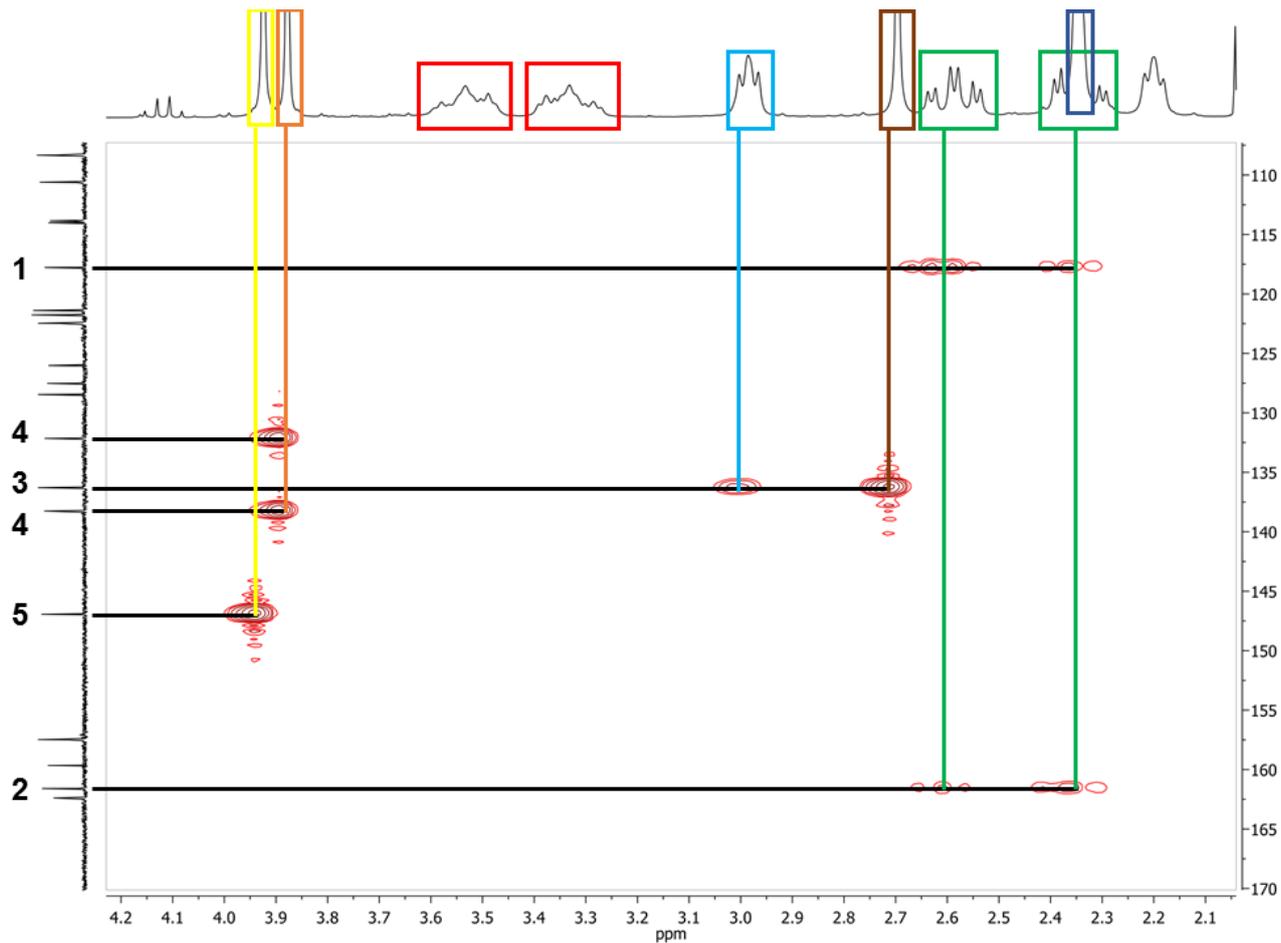
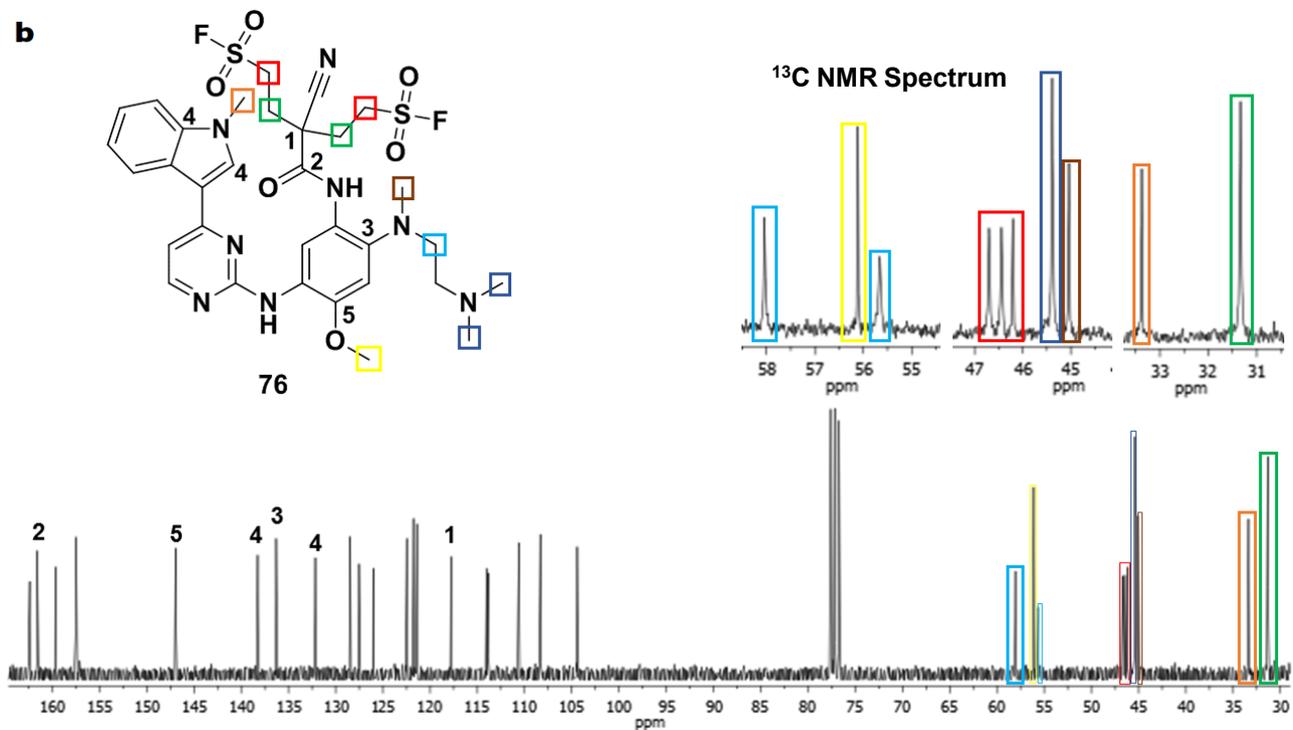
A selected portion of the HSQC NMR spectrum of compound **76** is shown on the previous page in **Figure 7.14b**. Working from this spectrum, we were able to correlate the protons of the solubilising group (dark blue, light blue, brown), aromatic methyl ether (yellow) and indole methyl (orange) to their appropriate carbons, which can be referenced to their respective colours in the structure displayed in **Figure 7.14a**. However, we were most interested in identifying the carbons of the proposed doubly substituted sulfonyl fluoride functionality. Agreeably, both sets of ethylene protons (green and red) suspected to belong to this moiety coupled to one carbon each. Examination of the most upfield of these sets revealed both of the triplet of doublet signals at δ 2.34 and 2.59 ppm (green) coupling to the carbon signal at δ 31.3 ppm. Comparison of the larger relative peak intensity of this carbon signal to the neighbouring indole methyl substituted carbon (orange) at δ 33.4 ppm reinforced our belief that this was two ethylene carbon signals instead of one.

Focussing on the more downfield set of multiplets (red) at δ 3.34 and 3.53 ppm, we again saw a correlation of both of these sets of protons to a single carbon signal at δ 46.4 ppm. Interestingly, this carbon signal appeared to be split into an apparent triplet. However, this was not unexpected, as the proposed carbon centres (red) were in close proximity to the fluorine atom of the sulfonyl fluoride moiety. As fluorine atoms are known to couple to neighbouring carbon nuclei due to the high abundance of the ^{19}F isotope and lack of C-F decoupling in standard ^{13}C NMR spectroscopy pulse sequences, C-F splitting is observed and has been found in the ^{13}C NMR spectra of several fluorine containing compounds synthesised in this thesis. Therefore, we believed that the observed triplet was the result of an overlap between the two anticipated doublets of the two respective carbon centres (red). The observed coupling to and splitting of this carbon centre provided further weight to our assignments and the proposed structure of final compound **76**.

To gain further insight into the structural elucidation and confirmation of compound **76**, we undertook an HMBC NMR spectroscopy experiment, with a truncated section of the afforded spectrum shown in **Figure 7.15a** on the next page. We were able to identify several new carbon centres through long range interactions with previously identified proton signals, which are displayed as numbers on the structure provided in **Figure 7.15b**. These included the aromatic carbons **3**, **4** and **5**, which were identified from coupling to the solubilising group, indole methyl and aromatic methyl ether respectively.

Once again, investigation of the long-range interactions of the sulfonyl fluoride ethylene bridge protons (green and red) was the main priority in our spectral analysis. Therefore, inspection of the upfield region showed the first of these ethylene proton sets (green) coupling to two different two carbon signals at δ 117.8 and 161.6 ppm. While not displayed but notable, these protons also coupled to the other ethylene bridge carbon (red) at δ 46.4 ppm which appeared as a triplet, verifying its proximity to this carbon centre. By logical deduction of the shielding/deshielding effects experienced by the suggested neighbouring carbon centres, we proposed that the most upfield of these carbons at δ 117.8 ppm belonged to the quaternary carbon centre (**1**) and assigned the most downfield coupled carbon signal at δ 161.6 ppm to the carbonyl carbon (**2**) of the amide moiety. These assignments fulfilled the criteria of our postulated structure and further substantiated formation of the doubly substituted final compound **76**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

a HMBC NMR Spectrum**b**Figure 7.15: a) HMBC and b) ^{13}C NMR spectra of final compound 76.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Moving downfield to the remaining set of ethylene bridge protons (red), we did not discern any coupling to carbons within this region. While we did expect coupling to the proposed quaternary carbon centre (**1**), long-range coupling interactions are often dictated by the coupling constants and local geometry across the molecule, known as the Karplus relationship.⁴⁴ The Karplus equation, which describes a correlation between the ³J-coupling constants and dihedral torsion angles in NMR spectroscopy, shows that the magnitude of these couplings are near zero when the dihedral angles are near 90°. We believed this to be the case for the long-range coupling between the ethylene protons (red) and the quaternary carbon centre (**1**), as this take place through an aliphatic chain. As observed previously, we did see coupling of the ethylene protons (red) to the carbon (green) of the neighbouring ethylene protons at δ 31.3 ppm.

Lastly, the ¹³C NMR spectrum of compound **76** is shown in **Figure 7.15b**, with all assignments discerned through our 2D NMR analysis displayed either by colour or number. With all carbons correctly accounted for, the [M+H]⁺ ion of 733.2407 provided by HRMS matching the calculated mass of 733.2402 and investigation by 2D NMR spectroscopy substantiating the proposed structure of compound **76**, we were able to confidently conclude that di-substitution had occurred and resulted in exclusive formation of compound **76**.

With structural confirmation of compound **76**, we considered the synthetic adjustments required to afford the desired mono-alkylated compound **75**. Multiple attempts with replacement of solvent, lowering of reaction temperature and alteration of the equivalents of potassium carbonate only afforded the di-substituted product **76** in lower yields than the optimal conditions of **Scheme 7.17**. Interestingly, utilisation of a 2-fold excess of the cyanoacetamide driving group **74** also resulted in exclusive synthesis of the di-substituted product and a return of the excess starting material. Observing only sole formation of the dialkylated product **76** in these experiments, we sought to explore the implications of exchanging the base.

Piperidine possesses a similar pK_aH (~ 11) to potassium carbonate and is customarily used to affect enol formation in 1,3-dicarbonyl compounds, particularly in the Knoevenagel Condensation reaction.^{45, 46} Finding these properties desirable, we elected to employ piperidine as base in the reaction of **74** and ESF (**20**). Unfortunately, use of a catalytic, equimolar and excess amount resulted in no reaction taking place and return of the starting material. As piperidine is a nucleophilic base, we ascribed the absence of any product formation to a pre-emptive reaction between piperidine and the highly reactive ESF electrophile. When used catalytically, the piperidine would be completely consumed before being able to facilitate deprotonation of **74**. Similarly, in a stoichiometric or excess amount, all the ESF would react with piperidine, resulting in non-availability of electrophile and the observed event of no reaction taking place.

As recourse, we altered the base to the sterically congested potassium *tert*-butoxide. Owing to its non-nucleophilicity and relatively high strength (pK_aH ~ 17), potassium *tert*-butoxide has found utility as a base catalyst in organic synthesis for nearly 70 years.⁴⁷ However, as witnessed in previous reactions with potassium carbonate, use of either catalytic or stoichiometric amounts of potassium *tert*-butoxide led to conversion of **74** into the di-substituted product **76**. This was observed even at low temperatures ranging between -40 °C and -78 °C, with the amount of starting material returned inversely proportional to the equivalents of electrophile **20** employed.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Frustrated by the lack of control in product outcome, we ventured toward full deprotonation of **74** prior to treatment with the electrophile **20**. Consequently, a suspension of excess sodium hydride in THF was treated with **74** at -40 °C, allowed to warm to room temperature and stirred for one hour. After recooling to -40 °C, this solution was treated with **20**, followed by immediate TLC analysis. Intriguingly, visualisation revealed a very faint secondary product spot in between the starting material **74** and an intense spot correlating to the di-alkylated product **76**, potentially indicating minor generation of the mono-substituted product. Disappointingly, all attempts to isolate and purify this spot resulted in a mixture with either **74** or **76**, or too little product to discern and confirm successful synthesis of the mono-substituted **75**.

Perusal of the literature revealed a single publication related to fluorosulfonylethylation of active methylene substrates. With results comparable to those obtained in our synthetic endeavours, Krutak and co-workers were also unable to monoalkylate various malonyl derivatives in the presence of triethylamine, as rapid formation of the bis(fluorosulfonylethyl) derivatives observed and exclusively obtained.⁴⁸ From these and our experimental observations and outcomes, we put forth a plausible reaction mechanism to explain selective di-alkylation over mono-alkylation. This proposed mechanism may be found on the following page, illustrated in **Figure 7.16** and will be reviewed comprehensively.

Upon treatment of **74** with an applicable base (green), we believe deprotonation of the acidic cyanoacetamide α -proton leads to formation of a carbanion. This carbanion is stabilised by the neighbouring electron-withdrawing groups, able to form resonance structures such as the enolate structure **77**. This nucleophile then reacts with the electrophile alkene of ESF (red) to form intermediate **78** in a conjugate addition reaction. Hitherto, this is the standard reaction mechanism for a Michael addition reaction. Traditionally, proton abstraction by the enolate intermediate of **78** (red, left) would occur from either the protonated base or solvent. Owing to its proximal location and acidity from the adjacent electron-withdrawing groups, one could propose that the remaining active methine proton (purple) could be abstracted by the neighbouring enolate functionality (**78**, red, left). However, such a shift would require an antarafacial [1,3] sigmatropic rearrangement. Although allowed, these rearrangements are virtually unknown and not observed, due mainly to the unfavourable orbital arrangement/asymmetry and consequent geometric requirements for the reaction.⁴⁹ Yet, this intermediate is able to invert to the anion stabilised resonance form of **78** (red, right). The tetrahedral character of the sulfonyl fluoride group and adjacent anion (red) allows for stabilisation of the neighbouring methine proton (purple) through induction. Therefore, we postulate that the enolate functionality could act as an intermolecular base, either reinitiating the reaction through deprotonation of another equivalent of **74** or abstracting an active methine proton (purple) from another formed resonance intermediate such as **78**. Additionally, abstraction from the protonated base may also occur, which may in turn deprotonate this active methylene proton (purple). Furthermore, this exchange may be driven by the inductive effect of the sulfonyl fluoride group, stabilising the newly formed anion. In doing so, a carbanion similar to that formed in the initial steps of the reaction could then be generated at the α -position of the malonyl derivative, which may in turn convert into the nucleophilic resonance structure shown in **79**. We propose that this nucleophile then attacks a second equivalent of ESF (blue), to produce intermediate **80**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

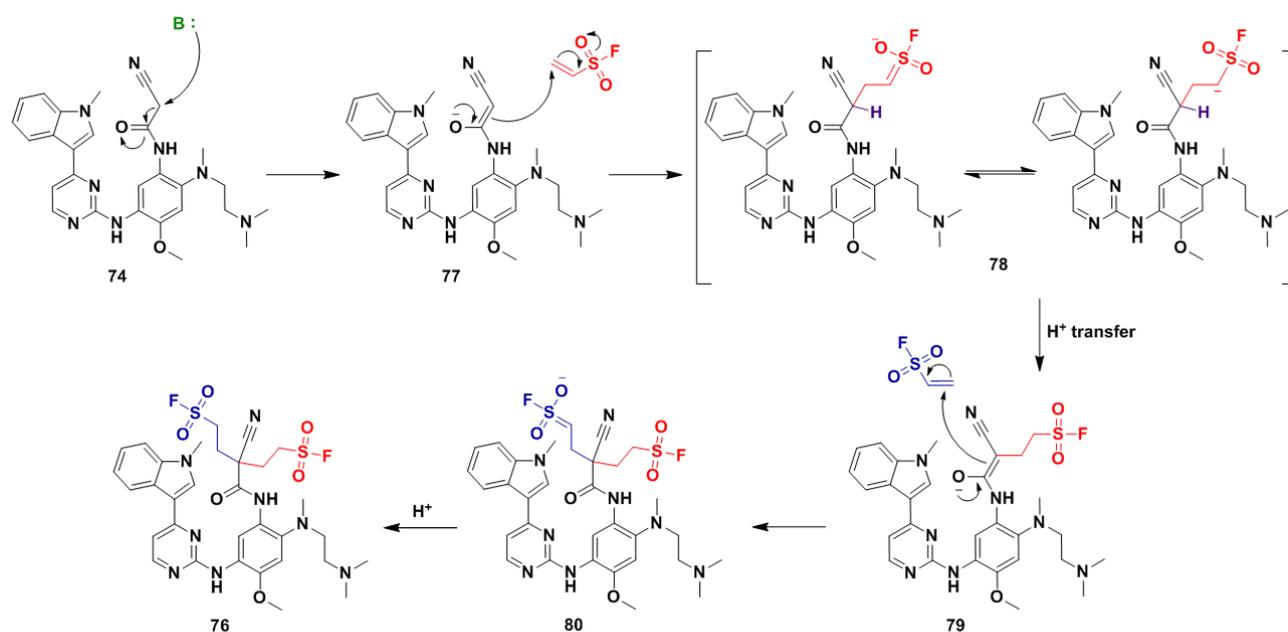
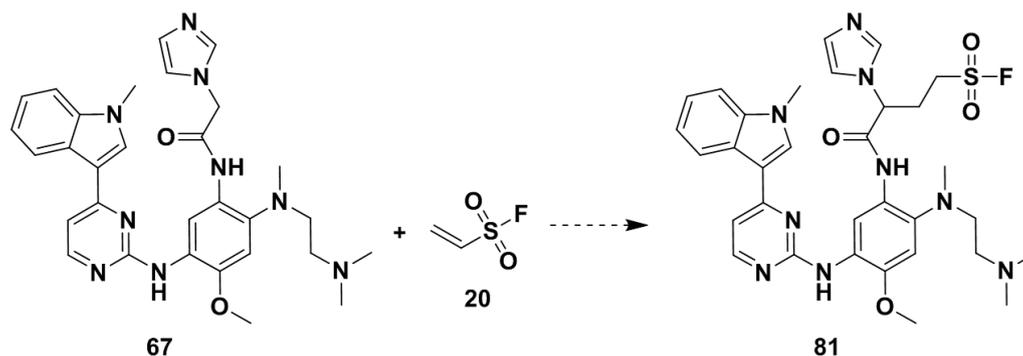


Figure 7.16: Proposed mechanism for the formation of compound **76**

Once again, the newly generated enolate functionality (blue) of **80** could potentially act as a base and deprotonate another equivalent of **74**, driving the reaction forward, abstract a proton from the external base or be ultimately quenched during the work-up procedure. All these possibilities lead to the generation of **76**, the major product isolated in all our reaction attempts.

Illustrated in **Scheme 7.18** on the next page, our final synthetic effort involved a Michael addition reaction between the imidazole-bearing compound **67** and ESF (**20**), in the hopes of obtaining the dual functional warhead **81**. Given the weaker electron-withdrawing capabilities of the adjacent imidazole group, we perceived the acidity of the active methylene proton to be reduced. We therefore reasoned that only potassium *tert*-butoxide and sodium hydride would be appropriate to facilitate deprotonation of **67**. To our disappointment, all attempts to synthesise final compound **81**, utilising these bases and parameters derived from the cyanoacetamide counterpart **74**, were met with no visible reaction taking place and a diminished return of starting material **67**. Due to time constraints, synthesis of both **75** and **81** could not be completed or investigated further. However, potential solutions to the problems encountered with this class of compound will be examined in the future work section (**Section 7.7**).

This concluded the synthesis of our osimertinib derived serine-targeting compounds, allowing us to proceed with biochemical and cellular evaluation of the successfully synthesised final compounds up until this point. These compounds were screened against the wild-type and various mutant forms of EGFR, with the intention of either reversibly interacting or covalently modifying the mutant Ser797 residue. An in-depth discussion of the results obtained from this assessment will be undertaken in the following section.

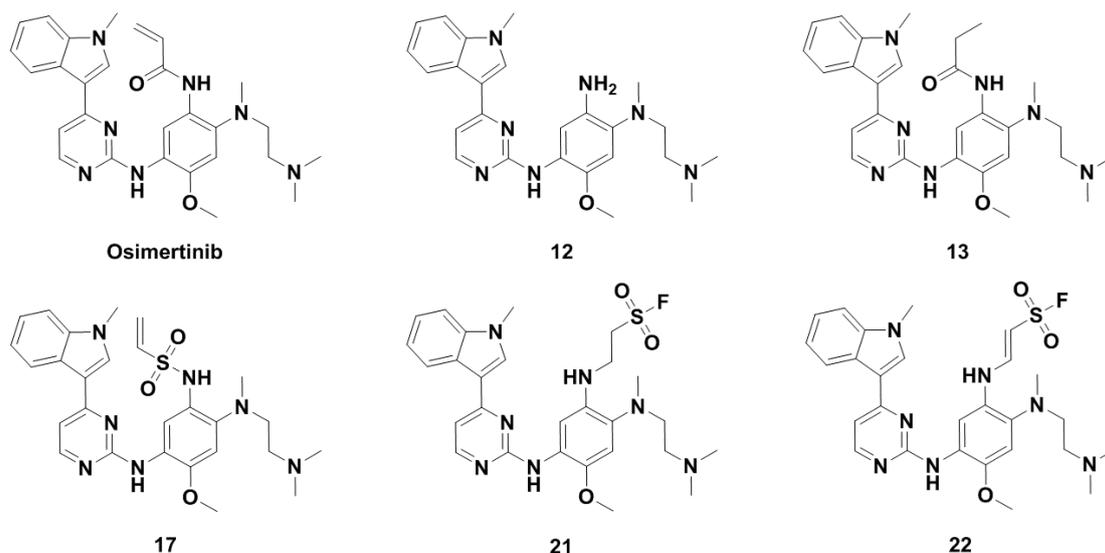
Scheme 7.18: Attempted synthesis of compound **81**.

7.5 Biochemical and Cellular Evaluation

As with the previous chapter, complications within the biology lab of our collaborators at the Technische Universität in Dortmund prevented biochemical inhibitory measurements of the synthesized final compounds against the triple L858R/T790M/C797S mutant variant of the enzyme. However, we were able to acquire IC_{50} values against the L858R/C797S double mutant, which could still provide valuable insight into potential interactions with the mutated Ser797 residue. Furthermore, a wild-type-bearing cell line A431, the EGFR-L858R/T790M drug resistant cell line H1975 and clinically relevant osimertinib resistant cell line EGFR-PC9/T790M/C797S was also employed. Additionally, promising compounds were further evaluated against BaF3 cell lines of both the single activating EGFR-L858R and the EGFR-T790M/C797S mutant variants of the enzyme (**Section 4.6.3**). Compound assessment was carried out either by Dr. Jonas Lategahn, Dr. Marina Keul or Mr. Tobias Grabe, with all measurements being executed in at least triplicate. The respective IC_{50} and EC_{50} values of osimertinib, the unsubstituted scaffold **12** and the reversible propionamide counterpart of osimertinib **13** accompany the obtained results in the report.

An overview of the results from the biochemical and cellular evaluation of the Michael acceptor- and sulfonyl fluoride-containing final compounds may be found on the following page in **Table 7.1**. Inspection of the first entries of the assessed inhibitors, the unsubstituted aniline scaffold **12** and propionamide analogue of osimertinib **13**, revealed an anticipated loss in activity when compared to osimertinib. This decrease was witnessed across-the-board, both biochemically and in a cellular setting, and can be ascribed to the loss of irreversible inhibition of the EGFR enzyme in the absence of the acrylamide Michael acceptor moiety. While compound **12** displayed a stronger biochemical profile than **13**, these results were reversed for the cellular assessment. This pointed towards an intolerance for the naked aniline within cells and a preference for substitution at this position with an appropriate functional group. Notably, the reduction in efficacy for **13**, when compared with osimertinib in the cellular measurement against the PC9/T790M/C797S triple mutant, was not as pronounced, with respective EC_{50} values of 3.9 and 3.1 μM . As we wished to investigate the potential interactions of our electrophilic fragments with the mutated Ser797 residue, our principal manner of comparison was with the unsubstituted aniline **12** and the propionamide-bearing compound **13**, and secondarily with osimertinib where applicable.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Table 7.1: Biochemical and cellular evaluation results for the Michael acceptor and sulfonyl fluoride containing synthesised final compounds.

Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]				
	wild-type	L858R	L858R/ T790M	L858R/ C797S	A431	H1975	PC9/ T790M/ C797S	BaF3/ L858R	BaF3 T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	261 ± 84.2	756 ± 340	16.0 ± 5.00	3109 ± 648	< 14	2041 ± 590
12	80.3 ± 23.1	10.5 ± 1.66	12.0 ± 1.04	-	7499 ± 2641	10260 ± 1783	5395 ± 1860	-	-
13	216 ± 58.7	33.8 ± 2.41	56.7 ± 4.29	-	8501 ± 620	8943 ± 1890	3937 ± 470	-	-
17	32.3 ± 13.6	7.86 ± 0.30	3.87 ± 0.47	154 ± 36.5	6265 ± 3245	4762 ± 1912	6341 ± 2847	< 14	24853 ± 8911
21	42.5 ± 12.0	6.93 ± 3.48	4.87 ± 1.04	693 ± 134	28159 ± 2611	30000 ± 0	24003 ± 3274	-	-
22	56.2 ± 27.1	5.64 ± 2.13	11.8 ± 2.11	92.9 ± 8.20	6913 ± 1461	10070 ± 220	5128 ± 561	-	-

Examination of the biochemical data of the sole Michael acceptor-containing final compound **17** exhibited an overall improved inhibitory profile in comparison to the reference compounds **12** and **13**. Whilst minor for the wild-type and single activating L858R mutation, this enhancement was a more than 3- and 14-fold improvement over **12** and **13** respectively for the L858R/T790M mutant, with an IC₅₀ value of 3.87 nM. This alluded to potential covalent modification of the mutant enzyme through a Michael addition reaction between the ethenesulfonamide moiety of **17** and the Cys797 residue. Notably, compound **17** also demonstrated a 1.7-fold improved IC₅₀ value against the L858R/C797S double mutant when compared to osimertinib. While this improvement was not indicative of irreversible binding, it implied that the more electron-deficient system of the ethenesulfonamide electrophile was more suitable for reversible interactions with the mutated serine residue than the traditional acrylamide electrophile.

As seen with the unsubstituted scaffold **12** and propionamide **13**, the EC₅₀ values of compound **17** drastically increased in the cellular assessment against the enzyme. This inflation to single digit micromolar values strongly supported a lack of covalent inhibition taking place within the cellular setting. Comparable to the biochemical results, compound **17** retained a 2-fold superiority over both model compounds **12** and **13** against the H1975 double mutant containing cell line.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

This highlighted the ethenesulfonamides greater affinity for reversible interaction with the Cys797 residue over the unsubstituted aniline (**12**) or propionamide (**13**) functionalities. However, this edge in activity was surprisingly not reflected in the EC_{50} values obtained against the PC9/T790M/C797S triple mutant, with compound **17** exhibiting an EC_{50} value of approximately $6.3 \mu\text{M}$ – weaker than both reference compounds **12** and **13**. This signified a 2-fold loss of activity compared to osimertinib against this mutant variant. These results suggest that this electrophile is more applicable for targeting cysteine residues and ill-suited for interactions with the mutated Ser797 residue in a cellular setting. Additionally, as osimertinib was specifically developed to contain an acrylamide Michael acceptor, we propose that the interchange to an ethenesulfonamide electrophile in compound **17** could have negatively affected the pharmacokinetic profile of the inhibitor. As such, we believe these factors to be responsible for the poor performance of this compound against the PC9/T790M/C797S.

Owing to the superior biochemical activity of **17** over osimertinib in inhibiting the clinically relevant L858R/C797S mutant, the ethenesulfonamide-containing compound was subsequently evaluated against the available Ba/F3 cell lines. Compound **17** matched the assay limit of 14 nM displayed by osimertinib against the single Ba/F3-L858R mutant but exhibited a dramatic loss in activity to nearly $25 \mu\text{M}$ against the double Ba/F3-T790M/C797S mutant. This result, combined with the PC9/T790M/C797S data, suggests that use of this electrophile will not be applicable in overcoming enzymes bearing the C797S mutation.

Moving to the sulfonyl fluoride-containing inhibitors **21** and **22**, we found an inhibitory profile similar to that of compound **17** against the wild-type, single and double mutant bearing enzymes. However, the biochemical data obtained for the key L858R/C797S double mutant variant was particularly compelling. Inspection of the entries for the saturated (**21**) and unsaturated (**22**) sulfonyl fluorides demonstrated IC_{50} values of 693 and 92 nM respectively. Impressively, this equated to a 3-fold improvement in inhibition for compound **22** when compared to osimertinib, and a 2.7-fold loss in activity for the structural counterpart **21**. Exploring the obtained values of the cellular assays, the saturated **21** displayed little to no activity for the wild-type and double mutant H1975 cell lines but showed improved efficacy against the PC9/T790M/C797S triple mutant. These results were mirrored for compound **22**, with the unsaturated analogue exhibiting an EC_{50} value of $\sim 5 \mu\text{M}$ against the triple mutant - a 2-fold improvement over the double mutant. Favourably, this observation indicated an interaction preference of these electrophiles for the mutated serine residue of the triple mutant over the typical cysteine residue present in the wild-type and double mutant variants of the enzyme.

We were most intrigued by the 7.5-fold difference in biochemical activity between the two sulfonyl fluorides, a trend which was preserved in the results of the cellular assays. The unsaturated sulfonyl fluoride **22** outperformed its saturated equivalent **21** by nearly 5-fold against the PC9/T790M/C797S cell line. Enthused by this variation in activity brought about by the presence of the adjacent double bond, we scoured the literature for a suitable explanation. From our search we uncovered a recent publication by Mukherjee et al. surrounding sulfonyl fluoride reactivity and hydrolytic stability.⁵⁰ In this study, the group investigated the reaction of various sulfonyl-fluoride containing warheads with nucleophilic *N*-acetyl amino-acid side chains of cysteine, tyrosine and lysine under physiological conditions. With emphasis on the influence of steric and electronic factors, numerous substituted aryl sulfonyl fluorides (activating and deactivating) and closely related vinyl and aliphatic sulfonyl fluorides were utilised.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

For the aryl sulfonyl fluorides, the group found that their reactivity towards the amino acids could be modulated accordingly by altering the electronic properties of the electrophile. Furthermore, the aryl sulfonyl fluorides stability towards hydrolysis was found to be inversely proportional to their displayed reactivity. However, these trends were not reflected in the highly reactive aliphatic and vinyl sulfonyl fluoride warheads. As these reactive groups are pertinent to our synthesised final compounds, we will focus solely on these results which can be seen in **Figure 7.17a**.

Phenylmethanesulfonyl fluoride (**PMSF**), pictured in **Figure 7.17a**, has found wide use in biological sciences as a protease/esterase inhibitor and covalent probe.⁵¹ Upon incubation with nucleophilic amino acid residues in a buffered aqueous solution of pH 7.5 at 37 °C, **PMSF** was found to rapidly hydrolyse. While in accordance with previous reports on the inhibitor, the group found that this decrease in hydrolytic stability did not correlate with an increase in reactivity, as was witnessed with the aryl sulfonyl fluoride electrophiles. By contrast, the rate of hydrolysis of the structurally related phenylethanesulfonyl fluoride (**PESF**, **Figure 7.17a**) was found to be approximately 13-fold slower than that of **PMSF**. However, despite this increased stability, the group did not observe sulfonylation of the nucleophilic tyrosine amino acid by **PESF**. Although less reactive than the aryl sulfonyl fluorides investigated, 2-phenylethanesulfonyl fluoride (**82**, **Figure 7.17a**) did undergo sulfonylation of tyrosine and was found to be a great deal more hydrolytically stable than **PMSF** and **PESF**. As a side note, no 1,4-addition of *N*-acetyltyrosine to the strongly electron-deficient olefin of the vinylsulfonyl fluoride was witnessed in any of the reactivity studies.

All the above observations are consistent with the proposed mechanism of hydrolysis for aliphatic sulfonyl fluorides, displayed in **Figure 7.17b**. Sulfonyl halides, including fluorides, which bear acidic protons in the α -position (red) have been documented to undergo an elimination reaction to form sulfene-type intermediates such as (**83**), with **PMSF** as prime example (**Figure 7.17b**).^{52, 53} Hydrolysis of this highly reactive intermediate then proceeds via an addition reaction, resulting in formation of compound **84**. This mechanism is supported by the low stability and reactivity displayed by **PMSF**, owing to the presence of the acidic benzylic α -proton (red).

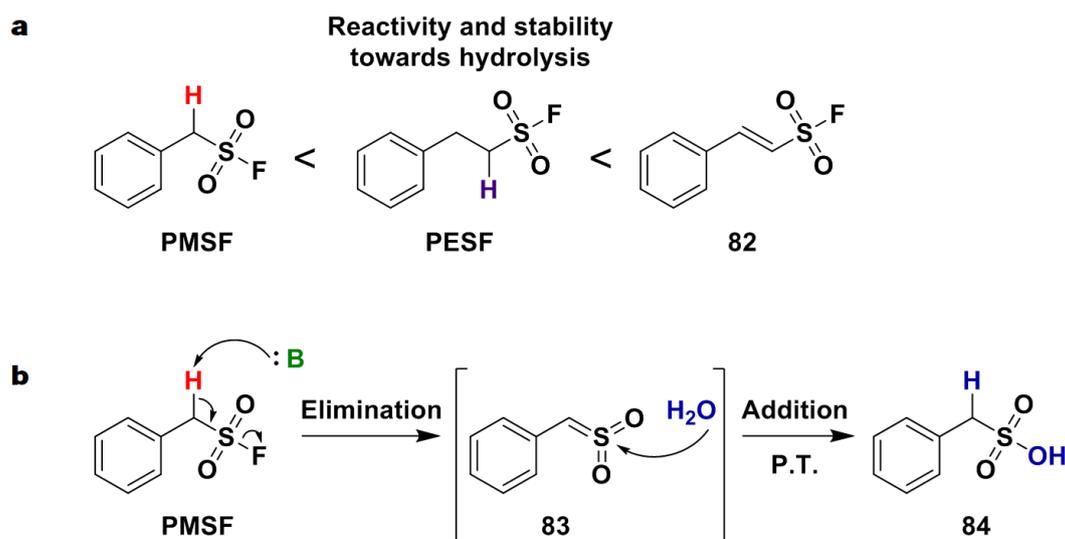


Figure 7.17: a) Depiction of the reactivity and hydrolytic stability trends displayed by the aliphatic and vinyl sulfonyl fluorides and b) Proposed mechanism of hydrolysis occurring through a sulfene-like intermediate **83** rather than via S_N2 reaction.⁵⁰

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

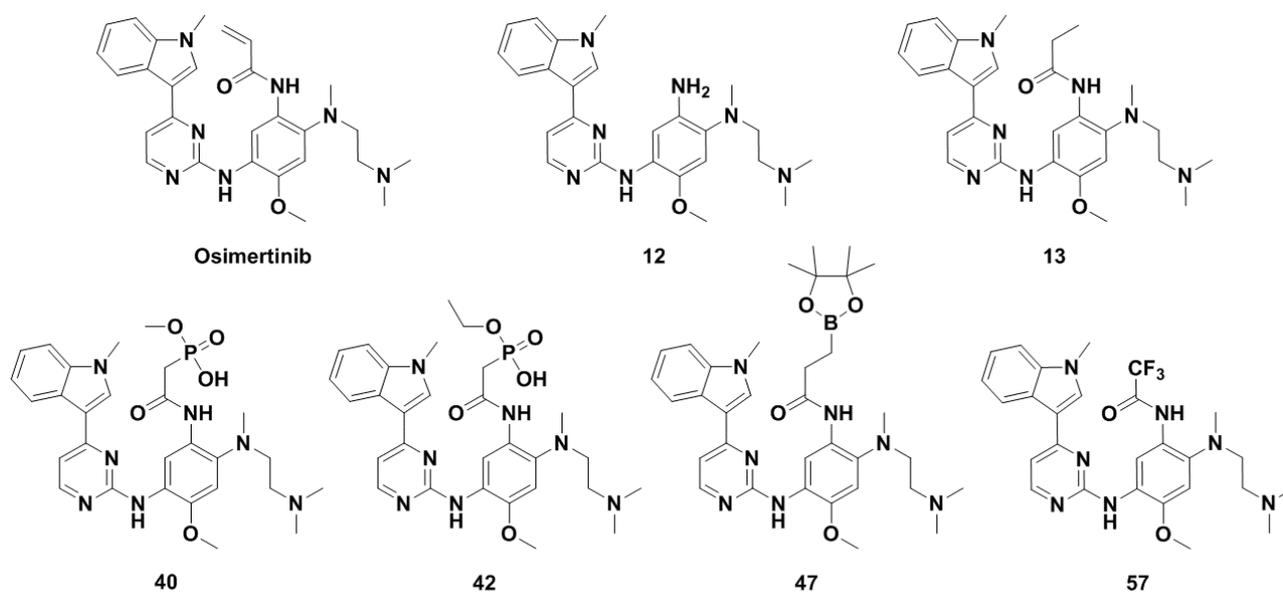
The greater stability exhibited by PESF can be attributed to the decreased acidity of the non-benzylic α -proton (purple), while its inert nature suggests that sulfene formation and subsequent hydrolysis occurs more rapidly than nucleophilic addition to the sulfonyl group by tyrosine.⁵¹ Importantly, the absence of an acidic proton in the α -position for the vinyl sulfonyl fluoride **82** resulted in a stable electrophile able to react with *N*-acetyltyrosine.

Considering these mechanistic insights, the disparity in inhibitory activity displayed by the aliphatic and unsaturated sulfonyl fluoride **21** and **22** respectively becomes transparent. We propose that the noticeably lower IC_{50} and EC_{50} values exhibited by the aliphatic sulfonyl fluoride **21**, compared to **22**, can be ascribed to circumstantial hydrolysis of the warhead, as the biochemical and cellular evaluation is carried out at pH 7.4. This is plausible, owing to the presence of an acidic α -proton, thereby allowing for hydrolysis via the discussed elimination/addition reaction sequence and formation of the sulfene-type intermediate. Hydrolysis by this mechanism is not possible for the unsaturated sulfonyl fluoride **22**, which does not contain an available acidic proton in the α -position. We believe it is this stability, and the ability of the sulfonyl fluoride warhead to survive and interact with mutated Ser797 residue, that gives rise to the greater inhibitory activity displayed by this compound. This is an interesting and positive result, as moderate reactivity and hydrolytic stability are favourable characteristics of electrophiles undergoing covalent modification. Therefore, it is our opinion that this class of warhead could see further use in the prospective development of new irreversible inhibitors which will be discussed in the future work section (**Section 7.7**).

The biochemical and cellular evaluation outcomes for the second set of compounds is illustrated in **Table 7.2**. Of this group, the first functionality we wished to scrutinize was the methyl and ethyl mono-phosphonate esters **40** and **42** respectively. As we were unable to isolate the highly reactive fluorophosphonate derivatives, covalent modification of Ser797 was ruled out. Therefore, we wished to investigate the potential for favourable reversible interactions between the phosphonate moiety and the mutant residue. Surveying the obtained values, we saw little difference in activity between the two inhibitors, with the smaller methyl counterpart exhibiting a slightly improved inhibitory profile. Biochemical evaluation of compound **40** against the L858R/C797S double mutant displayed a 10-fold drop in activity in comparison to osimertinib and the first indication of incompatible interactions between this functional group and the Ser797 residue.

As with all the compounds analysed thus far, a decrease in activity was observed in the cellular assessment of both **40** and **42**. However, this increase was significantly greater for these compounds, which resulted in high micromolar EC_{50} values across all cell lines. Believed to stem from a poor pharmacokinetic profile provoked by the phosphonic acid part of the mono-phosphonate ester, we wished to gain a better understanding of this cellular intolerance. Confirmation of this speculation was made evident in the assessment of the methyl mono-phosphonate ester **40** against the two Ba/F3 cell lines. Of all compounds screened against the Ba/F3-L858R cell line, compound **40** was the only inhibitor to display a dramatic loss of activity. With an EC_{50} value of 263 nM - nearly 19-fold greater than the assay limit of 14 nM which was achieved by all other candidates – it was apparent that this functional group induced a negative impact and was not tolerated in cells. This was further reflected in the results of the Ba/F3-PC9/T790M/C797S triple mutant, with an 8-fold loss of efficacy when compared to osimertinib. Thus, it was evident from these findings that use of this functional group will not be effective in creating favourable reversible interactions with the Ser797 residue and overcoming resistance mediated by the C797S mutation in EGFR inhibition.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Table 7.2: Biochemical and cellular evaluation results for the mono-phosphonate ester, boronic ester and trifluoromethyl ketone containing synthesised final compounds.

Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]				
	wild-type	L858R	L858R/ T790M	L858R/ C797S	A431	H1975	PC9/ T790M/ C797S	BaF3/ L858R	BaF3 T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	261 ± 84.2	756 ± 340	16.0 ± 5.00	3109 ± 648	< 14	2041 ± 590
12	80.3 ± 23.1	10.5 ± 1.66	12.0 ± 1.04	-	7499 ± 2641	10260 ± 1783	5395 ± 1860	-	-
13	216 ± 58.7	33.8 ± 2.41	56.7 ± 4.29	-	8501 ± 620	8943 ± 1890	3937 ± 470	-	-
40	306 ± 92.9	40.8 ± 3.14	70.5 ± 7.75	2515 ± 328	22001 ± 9333	29749 ± 369	10962 ± 410	263 ± 9.21	16645 ± 1171
42	334 ± 84.4	37.9 ± 5.30	76.1 ± 5.93	-	24689 ± 8598	30000 ± 0	12899 ± 1699	-	-
47	53.4 ± 11.4	10.9 ± 1.30	10.9 ± 0.46	1163 ± 283	13653 ± 2621	9631 ± 708	6863 ± 2270	< 14	5342 ± 374
57	6666 ± 1154	6966 ± 1356	20000 ± 0	-	19323 ± 3655	19600 ± 1467	30000 ± 0	-	-

Analysis of the inhibitory values of the boronic ester-containing compound **47** revealed some interesting features. Firstly, as this compound displayed an equivalent or improved biochemical profile over the unsubstituted aniline scaffold **12** and propionamide-bearing compound **13**, we concluded that the sizeable boronic ester was tolerated within the binding pocket. This was a pleasant surprise, as we expected the bulky pinacol group to create an unfavourable steric clash within the active site. Incubation of **47** with the L858R/C797S double mutant resulted in a ~ 4.5-fold loss of activity when compared to osimertinib, with an IC₅₀ value of 1.1 μM. While a moderate decrease, we found this result to be encouraging, as we maintain that the propenamide-linked boronic ester moiety is suitable for inhibition of mutant variants of the EGFR enzyme. Suitably, we anticipate that the hydrolysed boronic acid counterpart, while not afforded in this synthetic attempt, will be capable of surpassing compound **47** in its efficacy. This is due to its potential to form favourable reversible interactions with and irreversibly inhibit the mutant serine residue.^{14, 17}

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

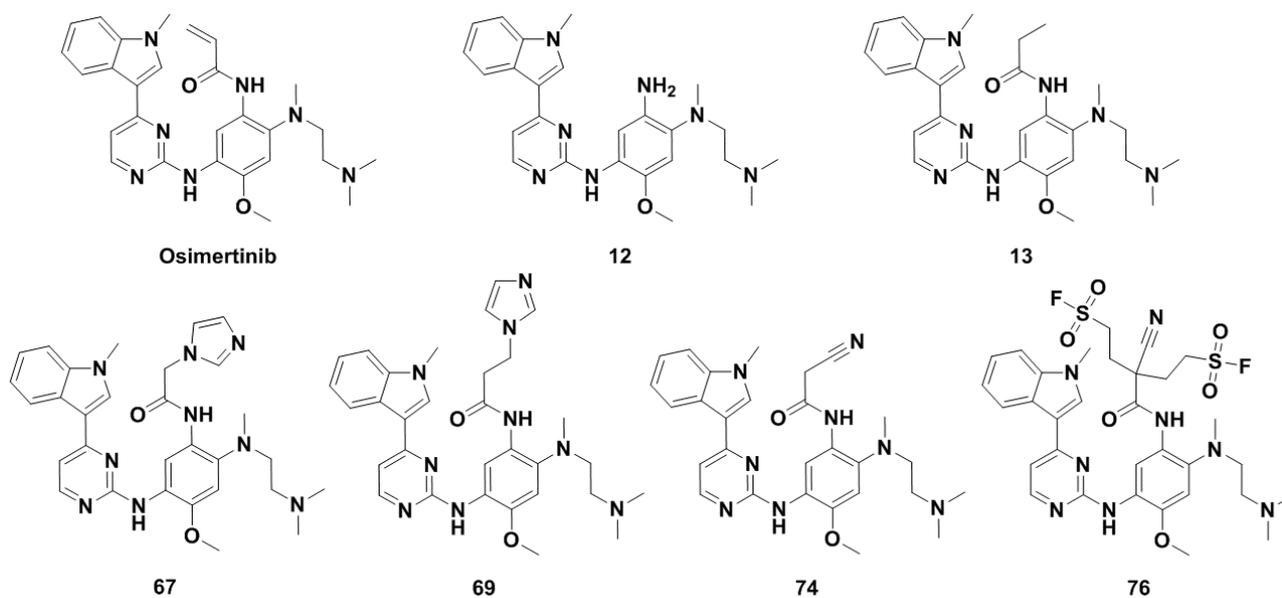
Inspection of the A431 wild-type and H1975 double mutant cell lines reveals a similar trend to that found in the biochemical results, with **47** exhibiting improved selectivity and EC_{50} values over the unsubstituted aniline scaffold **12**. This however was not reflected in the results of the PC9/T790M/C797S cellular assay, with the boronic ester faring worse than both reference compounds **12** and **13** and displaying a near 2-fold loss in efficacy when compared to osimertinib. However, the obtained value of $6.8 \mu\text{M}$ was an improvement on most of the compounds already measured and on par with the ethenesulfonamide and unsaturated sulfonyl fluoride **17** and **22** respectively (**Table 7.1**). Encouragingly, compound **47** showed improved cellular inhibition of Ba/F3-T790M/C797S over the PC9/T790M/C797S triple mutant, with an EC_{50} value of $5.3 \mu\text{M}$. This result contrasts with the measurements obtained for ethenesulfonamide **17** and the methyl mono-phosphonate ester **40**, which both sustained major losses of activity against this cell line. The exclusive enhanced efficacy exhibited by **47** against the Ba/F3-T790M/C797S cell line underlines the potential development and utilisation of boronic ester and acid moieties in overcoming C797S mediated resistance. Prospective employment of these functional groups and the synthesis thereof will be reviewed in the future work section (**Section 7.7**).

Unfortunately, all results obtained for the trifluoromethyl ketone-containing compound **57** were found to be lacklustre. High micromolar IC_{50} and EC_{50} values indicated intolerance for the electronegative character of the trifluoromethyl moiety. As with the mono-phosphonate esters, this functional group negatively impacted the pharmacokinetic profile of the inhibitor. Furthermore, the compound was found to be inactive against the PC9/T790M/C797S cell line, indicating incompatibility for reversible interactions between this group and the mutated Ser797 residue. As such, we decided not to deliberate further nor include this electrophile within future considerations.

The results of the final collection of compounds to be examined, tabulated in **Table 7.3**, included those containing an imidazole, nitrile and dual-warhead functionality. Beginning with the imidazole-bearing compounds **67** and **69**, we observed positive trends which reinforced the inclusion of this scarcely utilised moiety. In the biochemical evaluation against the single activating L858R and double L858R/T790M mutant, both compounds displayed adequate inhibition values. While more efficacious than propionamide **13**, both compounds were roughly 2-fold less active than the unsubstituted scaffold **12**. In these variants, the longer propenamide-substituted imidazole compound **69** exhibited marginally superior values. As witnessed with the large propenamide-substituted boronic ester **47**, this alludes to space within the binding pocket at this region capable of accommodating larger substituents. Against the L858R/C797S mutant, **67** and **69** displayed similar activity, with both compounds approximately 2-fold less potent than osimertinib and the smaller imidazole-bearing **67** edging over **69** with an IC_{50} value of 471 nM vs 578 nM respectively. We found these results acceptable, as these compounds were amongst the most effective inhibitors of this mutant variant when compared with all other synthesised final compounds.

As demonstrated in the biochemical evaluation, compound **69** proved marginally more efficacious than the shorter chain **67** and was exclusively selective for the double mutant H1975 cell line over the wild-type in the cellular assessment. While this trend was mirrored in the cellular assay, the full spectrum of EC_{50} values obtained for **67** and **69** were found to be superior to both the unsubstituted aniline scaffold **12** – the inverse of what was observed for the biochemical inhibition – and the more potent propionamide **13**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Table 7.3: Biochemical and cellular evaluation results for the imidazole, nitrile and dual-warhead containing synthesised final compounds.

Compound	EGFR HTRF IC ₅₀ [nM]				EGFR CTG EC ₅₀ [nM]				
	wild-type	L858R	L858R/ T790M	L858R/ C797S	A431	H1975	PC9/ T790M/ C797S	BaF3/ L858R	BaF3 T790M/ C797S
Osimertinib	0.81 ± 0.47	0.68 ± 0.48	0.55 ± 0.40	261 ± 84.2	756 ± 340	16.0 ± 5.00	3109 ± 648	< 14	2041 ± 590
12	80.3 ± 23.1	10.5 ± 1.66	12.0 ± 1.04	-	7499 ± 2641	10260 ± 1783	5395 ± 1860	-	-
13	216 ± 58.7	33.8 ± 2.41	56.7 ± 4.29	-	8501 ± 620	8943 ± 1890	3937 ± 470	-	-
67	184 ± 39.0	24.8 ± 1.31	27.8 ± 1.68	471 ± 235	6051 ± 2691	7150 ± 2678	3240 ± 367	< 14	3162 ± 587
69	118 ± 28.0	20.6 ± 1.11	15.6 ± 2.48	578 ± 72.5	5409 ± 2556	5779 ± 1082	3217 ± 651	-	4
74	127 ± 39.3	21.1 ± 2.23	21.7 ± 1.05	485 ± 40.7	8624 ± 2104	9005 ± 2396	3464 ± 479	< 14	4339 ± 595
76	750 ± 262	122 ± 16.3	183 ± 6.96	-	30000 ± 0	30000 ± 0	30000 ± 0	-	-

This result was significant for several reasons, the first being that this gave an indication of the imidazole moiety being well-tolerated within cells. Secondly, as compounds **67** and **69** were shown to be more effective than both reference compounds **12** and **13** against the double mutant H1975 cell line, we could conclude that the imidazole group was a more favourable substituent at this position. This outcome thereby implied that this functionality interacted more favourably with the Cys797 residue than the unsubstituted aniline **12** or propionamide **13**, as observed for ethenesulfonamide **17** (Table 7.1).

Even more compelling was the results obtained for the PC9/T790M/C797S triple mutant bearing cell line. Compounds **67** and **69** exhibited a near 2-fold improvement over their respective results from the H1975 assay, both with values of $\sim 3.2 \mu\text{M}$. Accordingly, this increased efficacy could be ascribed to the imidazole functional group interacting more favourably with the mutant Ser797 residue of the triple mutant than the Cys797 residue found in the H1975 cell line. Furthermore, these EC₅₀ values were comparable with that of osimertinib ($3.1 \mu\text{M}$).

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Enthused by these results, we assessed the potency of compound **67** against the available Ba/F3 cell lines, which reached the assay limit of 14 nM for the single activating mutant. Interestingly, with an EC₅₀ value of 3.1 μM, the imidazole compound **67** was found to retain and slightly improve its activity against the BaF3-T790M/C797S cell line when compared with the PC9/T790M/C797S triple mutant. This outcome, coupled with all other encouraging results afforded in the cellular assessment of compound **67** and **69**, highlights the prospect for development and employment of these imidazole functional groups in the therapeutic targeting of the EGFR-C797S mutant variant. We will review strategies for the incorporation and utilisation of this moiety within the future work section (**Section 7.7**).

Examination of the biochemical inhibitory profile of the cyanoacetamide-containing compound **74** reveals a striking similarity to that of the imidazole-bearing compound **67** and **69**. As such, our analysis of **74** will focus solely on the cellular evaluation and attempt to provide an explanation for the differences in activity observed between these functional groups in this setting. Accordingly, inspection of the values obtained in the H1975 cellular assessment of compound **74** reveals a major discrepancy when compared with the results of the imidazole-bearing compounds **67** and **69**. Compounds **67** and **69** were shown to be more efficacious than the reference compounds **12** and **13**, with EC₅₀ values of 7.1 and 5.8 μM respectively. In contrast, compound **74** exhibited a moderately weaker EC₅₀ value of 9 μM, comparable to that of the reference compounds. Intriguingly, evaluation against the PC9/T790M/C797S revealed a restoration of efficacy for the nitrile containing compound, with an exhibited EC₅₀ value of 3.4 μM - equivalent to both **67** and **69**. We postulate that this disparity in inhibition efficacy of compound **74**, which exists between both the Cys797 and Ser797 bearing enzymes and the imidazole-based inhibitors, may be explained on the premise of key differences in the structural and electronic elements of the interactive functional groups.

As both the imidazole and nitrile functional group contain nitrogen atoms that can serve as hydrogen bond acceptors, both classes of compounds are able to interact reversibly with the hard Ser797 hydroxyl functionality. This could provide an explanation for the improved activity witnessed for both the imidazole-bearing compounds **67** and **69** and the nitrile containing inhibitor **74**. Additionally, this gives clarity on why both imidazole compounds are nearly identical in potency and the nitrile slightly less so, proposed by observation to be a marginally weaker hydrogen bond acceptor.⁵⁴ This is further corroborated by the slightly higher obtained EC₅₀ value of 4.3 μM for **74** against the Ba/F3-L858R/C797S cell line, in comparison to the 3.1 μM displayed for imidazole **67**.

Although not favoured and leading to the observed drop in efficacy, both sets of compounds would have the nitrogen hydrogen bond acceptor at their disposal for interactions with the softer, more electron rich Cys797 residue. However, we believe the superior EC₅₀ values demonstrated by the imidazole compounds **67** and **69**, stems from the π-electron cloud of its heteroaromatic ring and rotational capabilities, which have been documented to participate in non-covalent interactions within protein structures.⁵⁵ Furthermore, sulphur-π interactions are prevalent in biochemistry and protein receptors, with cysteine/aromatic interactions playing important roles in protein folding and stabilisation and their presence being reported recently in solid state crystal structures of proteins.^{56, 57, 58}

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

We therefore propound that it is the presence of an additional sulphur- π interaction between the heteroaromatic π -electron cloud of the imidazole moiety and the Cys797 residue that is responsible for the enhanced activity displayed against the H1975 and A431 cell line in comparison to **74**. As the nitrile lacks this aromatic nature and the ability for sulphur- π interactions, we witness a decrease in activity to 9 μ M for compound **74** for the H1975 cell line. Notably, this loss of activity is also observed in the Cys797-containing wild-type variant of the enzyme.

Lastly, we propose that the longer propenamide-substituted imidazole **69** showed marginally superior activity over the shorter analogue **67** potentially due to a more favourable orientation and proximal location to the Cys797 residue. Importantly, the above hypotheses remain conjecture and cannot be confirmed until further studies are undertaken. This may include the use of molecular modelling or procurement of a crystal structure of either of the imidazole compounds complexed within any Cys797-containing form of the EGFR enzyme. Notably, as the nitrile exhibits a near 3-fold selectivity for the Ser797-bearing triple mutant over the Cys797 bearing enzymes, this electrophile could be developed and utilised in the selective inhibition of the EGFR-C797S mutant. This selectivity, combined with the other attractive inhibitory features discussed for the imidazole-bearing compounds, warrants further investigation of these functional groups in targeting the EGFR-C797S mutant.

The final inhibitor to be discussed for this chapter is the dual-warhead containing compound **76**. As the α -position of the cyanoacetamide was doubly derivatised with the saturated sulfonyl fluoride warhead, we anticipated a loss of activity due to the large size of **76** and the proclivity of the sulfonyl fluorides towards hydrolysis. The obtained results conformed to our expectations, with the dual-warhead displaying a dramatic decrease in biochemical inhibitory potency. Furthermore, the dual-warhead **76** was found to be inactive against all cell lines, presumably due to hydrolysis of the aliphatic sulfonyl fluoride moieties. This result would be in keeping with the poor EC₅₀ values measured for the sulfonyl fluoride inhibitor **21** (Table 7.1) and would almost certainly occur via the sulfene-type intermediate mechanism described previously (Figure 7.15b). Although this could be seen as a negative result, we believe that potential lies in the synthesis of the originally intended mono-substituted dual-warheads, particularly due to the interesting selectivity profile displayed by the nitrile compound **74**.

In conclusion, confirmation of the compounds binding properties and mode of inhibition was investigated using mass spectrometry experiments. Incubation of the synthesised final compounds with wild-type EGFR resulted in none of the incubated enzymes exhibiting an increase in mass equivalent to the corresponding compound labelled protein. This indicated a reversible binding mode for all inhibitors, which could be construed from the corresponding biochemical and cellular data of these compounds.

7.6 Conclusions

In this chapter, our research focus shifted from targeting the EGFR catalytic lysine residue to the development of inhibitors targeting the mutated Ser797 residue. As this mutant serine residue contains a harder, more inert hydroxyl nucleophile, we implemented a strategy of including electrophiles known to interact and irreversibly inhibit serine containing enzymes such as proteases and esterases. These electrophiles included Michael acceptors, sulfonyl fluorides, fluorophosphonates and phosphonate esters, boronic esters and acids, trifluoromethyl ketones, α -, β -haloketones, imidazoles and nitriles.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Furthermore, we envisaged the synthesis of dual-warhead-containing compounds, built on the premise of the catalytic triad in serine proteases. Building on the synthetic achievements of the previous chapter, these electrophiles would be incorporated on the osimertinib-derived scaffold at the aniline conventionally derivatised with the acrylamide Michael acceptor.

The synthesis of the osimertinib-derived driving group **12** proceeded smoothly, with high yielding intermediates contributing to an excellent overall yield and a multigram synthesis of the scaffold. This optimised synthetic pathway could allow for easy access in the future use of this heterocyclic driving group. For our electrophile synthesis and derivatisation, we were able to effectively synthesise the ethenesulfonamide-containing final compound **17** through spontaneous elimination of the sulfonylated intermediate. This was a noteworthy success, as we were unable to synthesise the lysine-targeting counterpart in the previous chapter. Similarly, oxidation of the aliphatic sulfonyl fluoride **21** to the unsaturated sulfonyl fluoride **22**, utilising manganese dioxide, was successfully carried out. To our knowledge, this is the first documented use of this class of electrophile in a potentially therapeutic context. This obscurity, combined with the encouraging inhibitory profile obtained, highlights the prospective development and employment of this rarely used functional group.

While unable to isolate the highly reactive fluorophosphonate final compounds, we were able to develop useful synthetic methodology and isolate valuable intermediates en route towards these inhibitors. This included the isolation of the carboxylic acid phosphonate ester fragments **33** and **34**, which could be employed in future large library syntheses through an amide coupling or Mitsunobu reaction with various driving group scaffolds. Additionally, the arduous identification and optimisation of a synthetic method for selective mono-hydrolysis of the ethyl phosphonate ester, using magnesium chloride as Lewis acid, may be of use in future synthetic endeavours.

Similarly, the hydrolysed boronic acid of compound **47** proved too sensitive for isolation. However, we were able to overcome difficulties generated by both the complexing/coordinative properties of the osimertinib solubilising group and hydrolysis of intermediate reactive fragments through employment of a *tert*-butyl ester functional group. Adaptation of our synthetic strategy rewarded us with the carboxylic acid boronic ester fragment **51**, which may be utilised in a similar manner described for fragments **33** and **34** in the generation of prospective compounds. Furthermore, there is no documented usage of this synthetic methodology to afford the amide boronic ester, underlining the novelty and potential of this approach.

The attempted synthesis of the α -, β -haloketone-containing compounds resulted in several interesting outcomes. Firstly, we were able to provide a potential explanation for the observed instability and inability to furnish this class of compound, both in this and the previous chapter. We proposed this mainly stemmed from interference by the basic osimertinib amine solubilising group. This was most clearly evinced in the attempted synthesis of the β -haloketone **65**, which underwent elimination to exclusively afford osimertinib (**14**). To our surprise, utilisation of α - and β -haloketone carboxylic acid fragments and CDI as amide coupling reagent resulted in the formation of the imidazole-bearing final compounds **67** and **69**. We have found no mention or use of this synthetic pathway within the literature, which provides a new and convenient one-pot method for the synthesis of this class of compound. Formation of compound **67** was confirmed by synthesis and coupling of the carboxylic acid imidazole fragment **72**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

While amide coupling with **72** has found application within the literature, our developed synthetic route towards this fragment, which makes use of the stable and easily removed *tert*-butyl ester group, has improved on all routes and yields towards this compound.

Lastly, synthesis of the envisioned dual-warhead-containing compounds proved difficult and ultimately unattainable. Employment of several bases in varying amounts, carried out in numerous solvents and at a range of temperatures all led to isolation of the doubly derivatised compound **76**. Spectroscopic analysis, using an array of techniques including 2D NMR spectroscopy, authenticated the successful synthesis of this compound. From the observations of our diverse experimental efforts, we proposed a mechanism for formation of compound **76**. Further strategies to obtain the desired mono-substituted dual-warhead will be disclosed within the following future work section.

Evaluation of the 14 synthesised final compounds revealed several compelling and encouraging results. While the ethenesulfonamide **17** exhibited a good biochemical inhibitory profile, particularly against the cysteine-bearing EGFR enzymes, this efficacy was not reflected in the cellular assays. Furthermore, this compound displayed a loss in efficacy against both the PC9/T790M/CYS797 and Ba/F3-T790M/C797S cell lines. As both these enzymes bear the mutant Ser797 residue, we concluded that use of this electrophile would not be suitable in overcoming resistance mediated by the EGFR-C797S mutation.

We were able to rationalise the disparity in activity between the aliphatic and unsaturated sulfonyl fluorides **21** and **22** respectively, through an investigation into the proposed hydrolysis pathway of these electrophiles. This mechanism, which proceeds via a sulfene-type intermediate, was shown to be exclusively available to the aliphatic sulfonyl fluoride **21**. This provided an explanation for the greater hydrolytic stability and consequent potency exhibited by the unsaturated counterpart **22**. This attractive feature, coupled with the previously mentioned novelty of this electrophile, emphasises its prospective use in medicinal chemistry. This will be discussed further in the following future work section.

The unremarkable activity displayed by the mono-hydrolysed phosphonate esters **40** and **41**, and the trifluoromethyl ketone-containing compound **57**, was attributed to the negative impact of these functional groups on the pharmacokinetic profile of the structure. As such, these compounds were not tolerated within a cellular context and we do not promote their further inclusion in targeting the EGFR-C797S mutant enzyme.

Contrary to our expectations, the data acquired for the bulky boronic ester compound **47** was found to be moderate but promising. We believe that successful hydrolysis of this moiety to the corresponding boronic acid, and the synthesis of similar boronic ester/acid warheads, could prove applicable and advantageous in overcoming the clinically relevant EGFR-C797S mutant variant. These boronic esters/acids, and the respective synthetic procedures required to access them, will be comprehensively discussed in the future work section.

Evaluation of the imidazole-bearing compounds **67** and **69** revealed positive results and interesting comparisons. The compounds performed well overall and were notably found to be the most efficacious inhibitors of both the mutant Ser797 residue containing cell lines PC9/T790M/CYS797 and Ba/F3-T790M/C797S. This was ascribed to interactions between the hydrogen bond acceptor nitrogen of the imidazole ring and the target serine residue.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

As the cyanoacetamide compound **74** contains a terminal nitrogen with similar hydrogen bond acceptor capabilities, the inhibitor displayed comparable activity to **67** and **69** against the C797S mutant variants of the enzyme. However, the EC₅₀ values deteriorated against enzymes bearing a cysteine residue at the 797 position. To rationalise this disparity in activity, we compared the structural and electronic elements of these functional groups. Accordingly, we attributed the enhanced activity of **67** and **69** to the aromatic nature of the imidazole ring and its ability to form sulphur- π interactions with the cysteine residue - a capability which the nitrile moiety lacks. Consequently, we see potential in the development of inhibitors containing either of these functional groups – the imidazole for its good overall interactive capabilities and the nitrile as a selective inhibitor of enzymes bearing the C797S mutation. Prospective examples of this may be found in the ensuing future work section.

Lastly, owing to the poor biochemical and cellular results exhibited, the doubly derivatised dual-warhead inhibitor **76** was presumed to have hydrolysed in a similar manner to that of the aliphatic sulfonyl fluoride **21**. While disappointing, we maintain that the synthesis of the initially envisioned dual-warheads, as well as other prospective electrophiles built on this principal, may serve as a worthwhile platform for investigation. The synthesis and application of these dual-warheads will be discussed in the future work section.

This brings us to the end of this chapter and concludes the research efforts of this PhD thesis. We will provide comprehensive concluding remarks on the project as a collective, future prospects for this research avenue and the outlook of EGFR inhibition and its clinical relevance in cancer treatment.

7.7 Future Work

For this section, we wish to showcase the broad applicability of incorporating the electrophiles and functional groups utilised within this chapter, as well as introducing new analogues of these functionalities. Based on their performance and potential for further development, we will focus exclusively on boronic esters/acids, sulfonyl fluorides, imidazoles and dual-warheads. As such, we will identify an appropriate example scaffold and outline the synthesis of our chosen functional groups and the requirements for their installation. All yields shown in schemes within this section are those provided within the relevant literature procedures.

As osimertinib was precisely designed to overcome resistance mediated by the EGFR-L858R/T790M double mutant, and not the EGFR-C797S variant, we will need to identify a driving group devised specifically for the triple mutant bearing enzyme. Furthermore, no crystal structure of osimertinib bound to this mutated enzyme exists within the literature or on the PDB. Therefore, for future considerations and development of this work, we would require a scaffold that fulfils the above-mentioned criteria. Compound **85**, shown in **Figure 7.18a**, was first mentioned in **Section 2.4.4 (Figure 2.15b)** as a potent inhibitor of the EGFR-del19/T790M/C797S triple mutant variant.⁵⁹ The *de novo* designed inhibitor utilises an unconventional phenolic hydroxy as hinge-binding motif and interacts reversibly with the mutant Ser797 residue through two hydrogen-bond interactions with the aminomethylpyrazine group (**Figure 7.18a**, orange). This resulted in **85** displaying an unprecedented 1000-fold selectivity for the triple mutant over the wild type, with an IC₅₀ value of 18 nM against the clinically relevant EGFR-del19/T790M/C797S mutant enzyme. This observed efficacy and selectivity, combined with the relative simplicity of the scaffold, leads to its selection as a prime candidate scaffold for inclusion and development of our electrophiles and functional groups.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

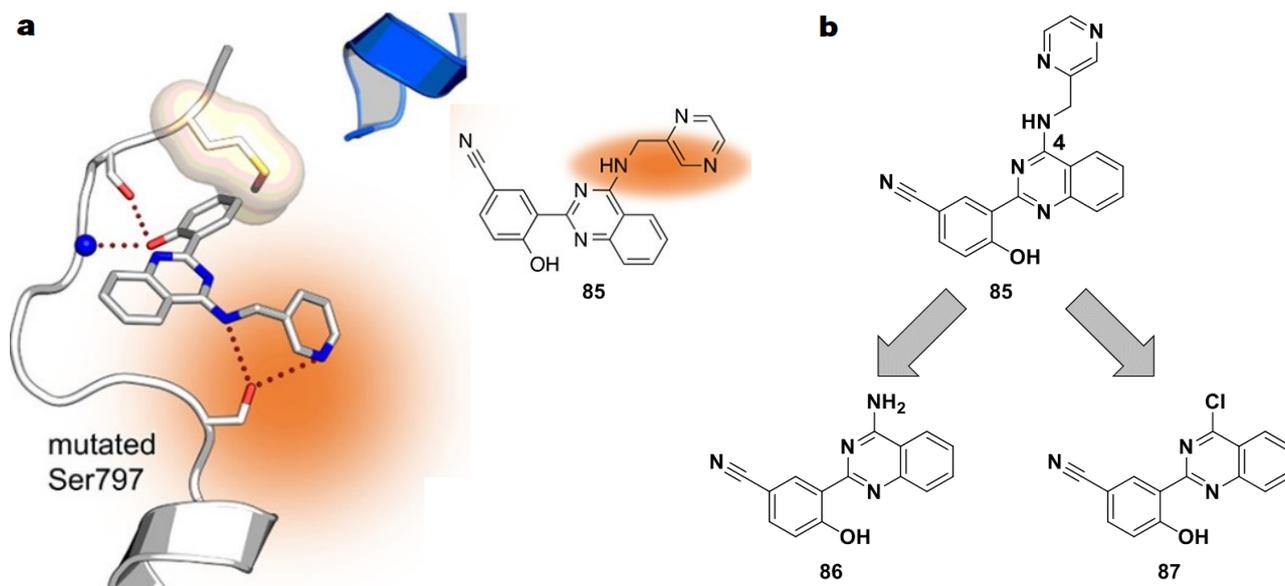
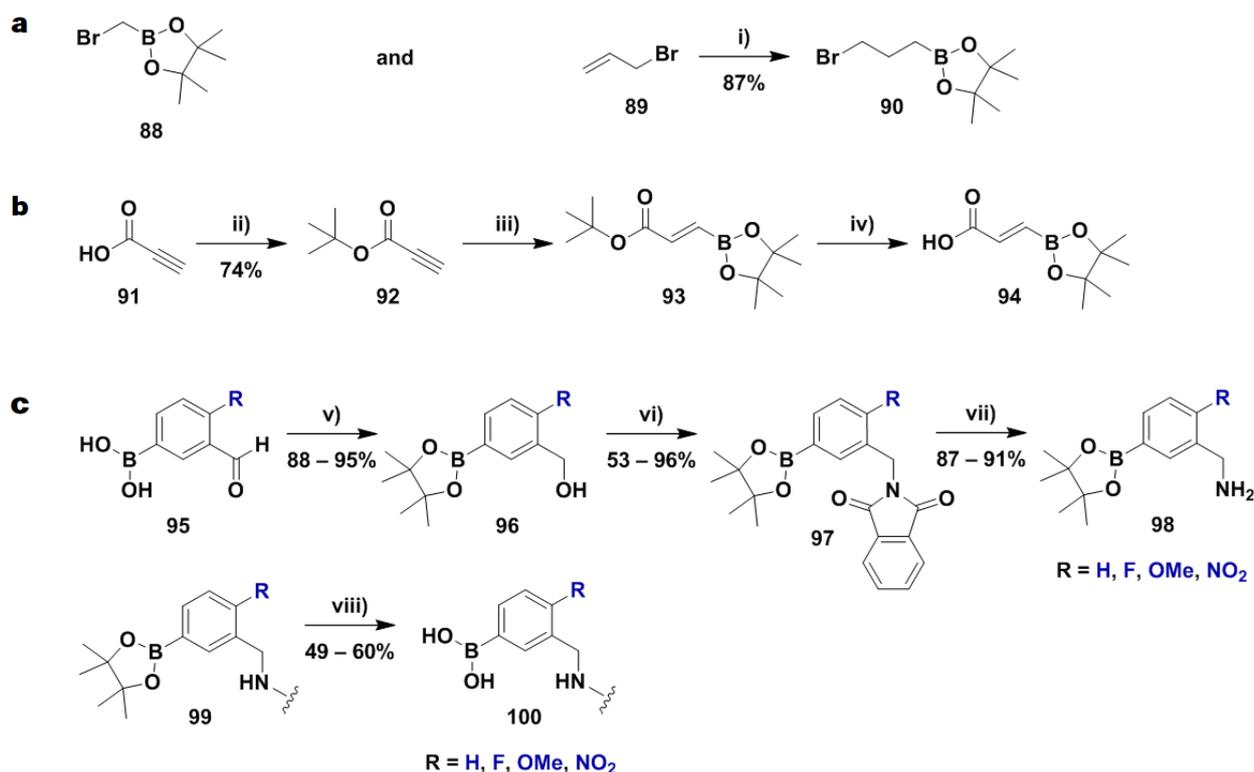


Figure 7.18: a) Binding mode of compound **85**, with novel hinge motif and two hydrogen bond interactions with the mutated Ser797 residue and b) Proposed heterocyclic scaffolds **86** and **87** to be utilised for derivatisation. Figure a) reproduced from Grabe et al.^{59, 60}

We envisage incorporation of these functional groups to occur at the 4-position of the quinazoline ring (**Figure 7.18b**), thereby replacing the aminomethylpyrazine. This will ensure a proximal location to the Ser797 residue, creating a high probability for formation of reversible interactions or potential covalent modification of the mutant serine. To do this, we propose synthesis of the 4-amino- and 4-chloroquinazolinone driving groups **86** and **87** respectively (**Figure 7.18b**). In doing so, we will allow for derivatisation at the designated 4-position with our chosen electrophiles and functional groups, in a similar manner to the heterocyclic scaffolds synthesised in **Section 5.4.1**.

The first functional group we will discuss is the boronic ester/acid, with the relevant prospective syntheses illustrated in **Scheme 7.19**. We wish to include and utilise two new and easily obtained aliphatic boronic esters, which are the commercially available bromomethane pinacol boronic ester **88** and the propyl analogue **90** (**Scheme 7.19a**). Boronic ester **90** may be afforded through triethylsilane mediated boration of allyl bromide (**89**) using boron trichloride at low temperatures.⁶¹ Owing to the successful use of a *tert*-butyl ester moiety in affording the carboxylic acid pinacol boronic ester fragment **51** (**Scheme 7.10**), we wished to repeat this route using propionic acid (**91**) to afford the unsaturated boronic ester **94**, as shown in **Scheme 7.19b**. We find the proposed synthesis of compound **94** to be intriguing, as this fragment contains two electrophilic sites – namely the Michael acceptor and boronic ester/acid. Therefore, this fragment could be used in the potential synthesis of dual-warhead inhibitors. The route towards **94** will begin with Steglich esterification of propionic acid **91** using DCC and DMAP as additive to afford propionic ester **92**.⁶² Following this, pinacol boronic ester **93** could be furnished through utilisation of the optimised parameters for copper catalysed β -boration using potassium *tert*-butoxide as base in a tetrahydrofuran/methanol mixture.³⁷ Subsequent hydrolysis of **93** could then be accomplished by stirring in a DCM/trifluoroacetic acid mixture at room temperature. This will afford the carboxylic acid fragment **94**, which may be further employed in amide coupling reactions.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



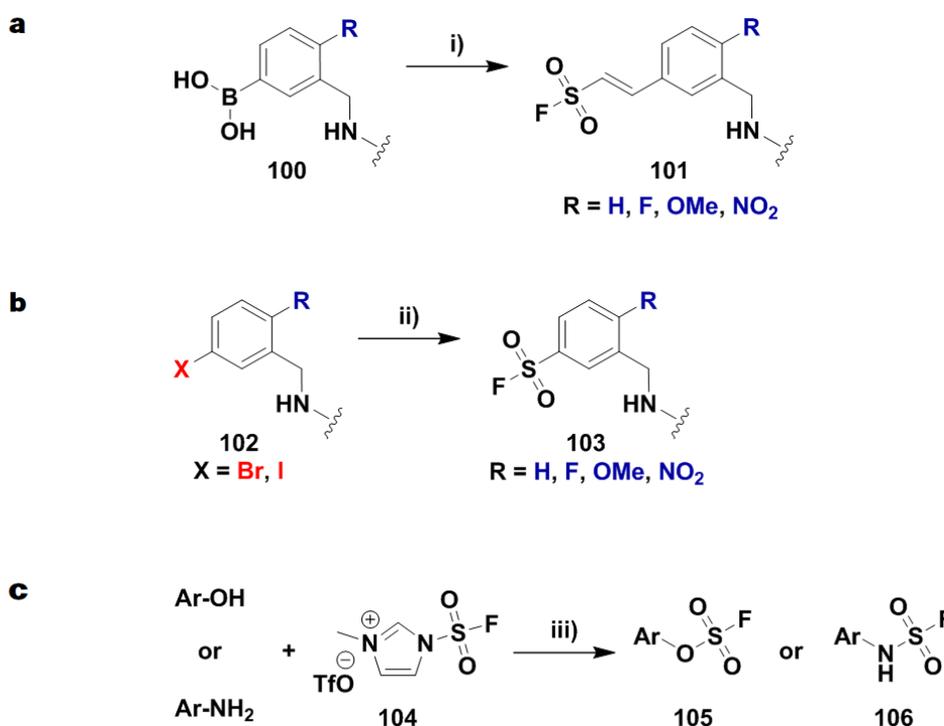
Scheme 7.19: Structures and proposed synthesis of a) Aliphatic boronic esters **88** and **90**, b) The unsaturated carboxylic acid boronic ester fragment **94** and c) Various aryl boronic esters with activating or deactivating substituents and the hydrolysis thereof. Reagents and conditions: i) triethylsilane (1.05 equiv.), BCl₃ (1.1 equiv.), pinacol (1 equiv.), Et₂O, -78 °C – rt; ii) *tert*-butanol (1 equiv.), DCC (1 equiv.), DMAP (10 mol%), DCM, 0 °C – rt; iii) bis(pinacolato)diboron (1.05 equiv.), CuCl (5 mol%), potassium *tert*-butoxide (10 mol%), xantphos (5 mol%), THF/MeOH (20:1), rt; iv) DCM/TFA (3:1), rt; v) pinacol (1.2 equiv.), MgSO₄ (2 equiv.), MeOH, rt then NaBH₄ (0.5 equiv.), MeOH, rt; vi) MsCl (1.5 equiv.), DIPEA (2 equiv.), DCM, 0 °C – rt then phthalimide (1.5 equiv.), K₂CO₃ (3 equiv.), DMF, rt; vii) hydrazine hydrate (3 equiv.), THF, 70 °C; viii) KHF₂ (6 equiv.), MeOH/H₂O (1:1), rt then SiO₂ (1.1 equiv.) EtOAc/H₂O (1:1), rt.^{61, 62, 63}

Pan and co-workers were able to develop a synthetic method towards boron-containing primary amines such as compound **98**, pictured in **Scheme 7.19c**.⁶³ The route, which begins from the appropriately substituted (3-formylphenyl)boronic acids **95**, proceeds via a Gabriel synthesis. This class of compounds piqued our interest, as these fragments would be able to mimic the structural elements of the aminomethylpyrazine of compound **85**, while incorporating a boronic ester/acid warhead. Furthermore, the inclusion of an activating or deactivating substituent at the *para*-position of the boronic acid would allow for an informative investigation into the effects of these groups on the reactivity of the warhead. These effects would be reflected in the biochemical and cellular evaluation of these compounds, evident as differences in efficacy. Furthermore, the substituted (3-formylphenyl) boronic acids **95** are commercially available, allowing for a favourable start to the synthesis. The first steps include boronic acid esterification with pinacol and reduction of the benzaldehyde to afford the corresponding alcohols **96**. This is followed by a Gabriel synthesis using phthalimide and hydrazine hydrate to facilitate aminolysis and furnish the desired boronic-ester containing benzylamines **98**. Lastly, the group were able to develop a mild hydrolysis procedure for the pinacol boronic esters using potassium bifluoride and silicon dioxide.⁶³ Following the subsequent coupling of **98** to our heterocyclic driving group **87** (**Figure 7.18**), this methodology could be utilised to hydrolyse the boronic esters of **99**, furnishing the desired boronic acid inhibitors **100**.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Opportunely, with the hydrolysed boronic acids (**100**) coupled to our driving group scaffold in hand, we will be able to employ an oxidative boron-Heck coupling developed by the research group of Arvidsson.⁶⁴ This will allow easy access to the vinyl sulfonyl fluorides **101** shown in **Scheme 7.20a**. This operationally simple, ligand- and additive-free reaction is compatible with the frequently used ethenesulfonyl fluoride and proceeds at room temperature using palladium(II)acetate, copper(II)acetate as oxidant and lithium acetate as base. As observed by the Grimster group in their study on sulfonyl fluoride reactivity, which was outlined in **Figure 7.17a**, we anticipate the vinyl sulfonyl fluorides (**101**) to be a great deal more stable than their aliphatic counterparts.⁵⁰ This can be mainly ascribed to the absence of an acidic proton at the α -position of the reactive group. Furthermore, as with the boronic acids, we will be able to explore the effects of the *para*-substituted electron-donating/withdrawing substituents with regards to the electrophile's interactions with the mutant Ser797 residue.

Due to their enhanced propensity towards hydrolysis, we wish to avoid the future use of aliphatic sulfonyl fluorides. Therefore, we envisage the synthesis of the stable aromatic sulfonyl fluorides **103** through a one-pot palladium-catalysed synthesis from the corresponding aryl halides **102** (**Scheme 7.20b**). Treatment of either a bromo- or iodo-equivalent aryl halide with the bis(sulphur dioxide) adduct 1,4-diazabicyclo[2.2.2]octane (DABSO) as sulphur dioxide source and Selectfluor as source of electrophilic fluorine has been documented to afford the matching aromatic sulfonyl fluorides in good yield.^{65, 66} Application of these reaction conditions could afford the aromatic sulfonyl fluorides **103**, of which the reactivity towards the mutated Ser797 residue could be compared through the various *para*-substituted activating and deactivating groups.



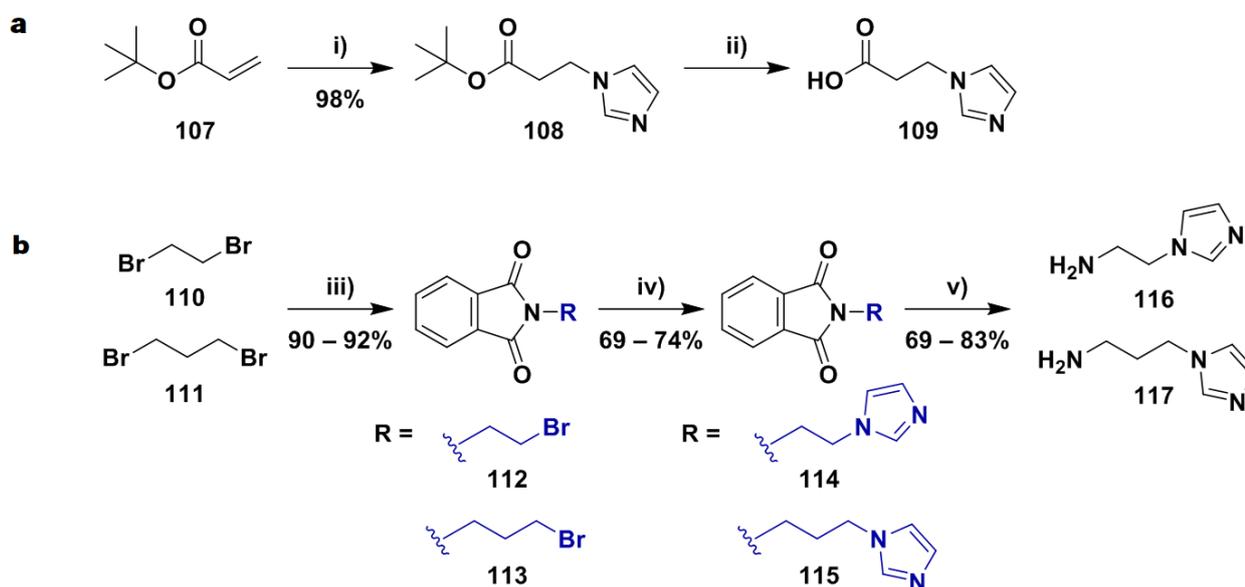
Scheme 7.20: Proposed synthesis of a) Vinyl aromatic sulfonyl fluorides **101**, b) Aromatic sulfonyl fluorides **103** and c) Fluorosulfates **105** and sulfamoyl fluorides **106**. Reagents and conditions: i) ESF (3 equiv.), Pd(OAc)₂ (10 mol%), LiOAc (1.2 equiv.), Cu(OAc)₂ (2 equiv.), THF, rt; ii) Pd(OAc)₂ (5 mol%), DABSO (1.2 equiv.), PdAd₂Bu (8 mol%), Et₃N (3 equiv.), *i*-PrOH, 75 °C then Selectfluor (2 equiv.), MeCN, rt; iii) **104** (1.2 equiv.), Et₃N (1.5 equiv.), MeCN, rt.^{64, 65, 66, 67}

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

In pioneering work from the Sharpless laboratory, synthesis of the fluorosulfonyl imidazolium salt reagent **104** allowed the group to transform phenols and anilines to their corresponding fluorosulfates **105** and sulfamoyl fluorides **106** respectively (**Scheme 7.20c**).⁶⁷ We believe this innovative synthetic protocol may be used to our advantage in the synthesis of warheads targeting and potentially covalently modifying the mutated Ser797 residue.

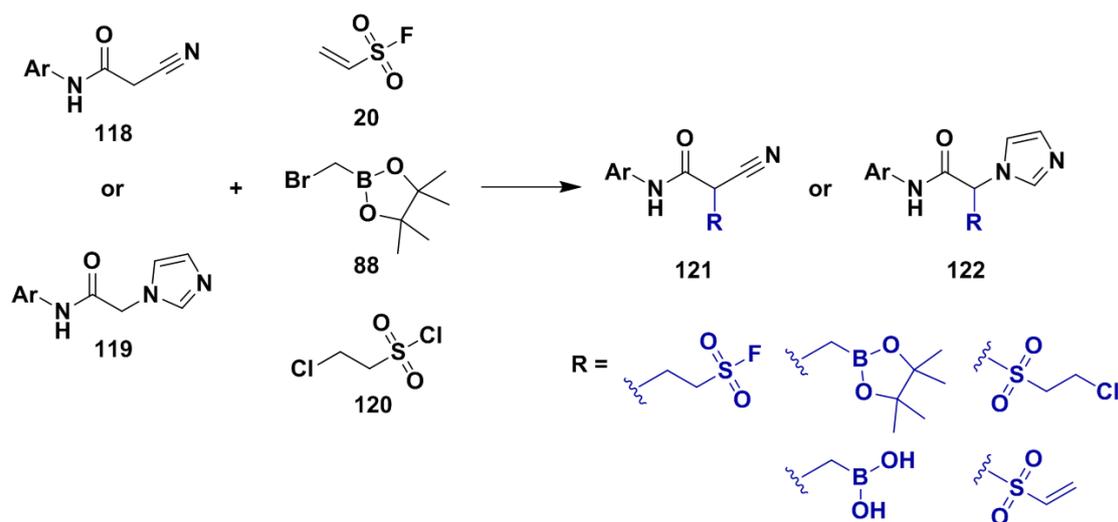
To complement the imidazole-containing carboxylic acid fragment **72** (**Scheme 7.15a**) synthesised in the last chapter, we wish to furnish the propyl analogue **109**, shown in **Scheme 7.21a**. As exhibited in the publication by Hou et al., this may be accomplished through a mild and efficient Michael addition of imidazole to the α , β -unsaturated *tert*-butyl acrylate (**52**) utilising potassium phosphate as base, thereby affording the *tert*-butyl ester imidazole **108**.⁶⁸ Subsequent hydrolysis of this ester, mediated by trifluoroacetic acid, will then provide the desired fragment **109**.

By including the amino-imidazoles **116** and **117**, found in **Scheme 7.21b**, a comparison of the reversible interaction capabilities of these moieties and that of the original aminomethylpyrazine used in compound **85** could be undertaken. Undergoing a Gabriel synthesis, dibromoethane or dibromopropane (**110** and **111** respectively) may be reacted with phthalimide to afford the intermediate derivatives **112** and **113**.⁶⁹ This will be followed by substitution of the alkyl halide with imidazole and aminolysis brought about by hydrazine hydrate whilst heating under reflux.⁷⁰ This will afford the amino-imidazole fragments, which may be coupled with the heterocyclic driving group scaffold **87**.



Scheme 7.21: Proposed synthesis of a) Carboxylic acid imidazole fragment **108** and b) Ethyl and propyl linked amino-imidazoles **116** and **117** respectively. Reagents and conditions: i) imidazole (1 equiv.), **107** (1.2 equiv.), K₃PO₄ (25 mol%), MeCN, rt; ii) DCM/TFA (3:1), rt; iii) phthalimide (1 equiv.), **110** or **111** (3 equiv.), K₂CO₃ (3 equiv.), TBAB (10 mol%), rt; iv) imidazole (1.2 equiv.), K₂CO₃ (2 equiv.), MeCN, 85 °C; v) hydrazine hydrate (3 equiv.), aqueous EtOH, 80 °C.^{68, 69, 70}

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



Scheme 7.22: Proposed synthesis of dual-warheads **121** and **122**.

Finally, we wish to discuss the potential development of the dual-warhead-containing inhibitors **121** and **122**, illustrated in **Scheme 7.22**. Due to the time constraints of this research venture, we were unable to effectively investigate mono-substitution at the α -position of cyanoacetamide **118** and the imidazole acetamide **119**. We maintain that with sufficient time and experimental effort, this novel synthetic transformation may be achieved and optimised. In addition to the originally intended use of ethenesulfonyl fluoride (**20**), we may attempt reactions with other electrophiles such as the commercially available bromomethane boronic ester **88** and chloroethane sulfonyl chloride **120**. As per our mechanistic investigation in **Section 7.4.3.9**, we propose that these reactions will not experience the same pitfalls associated with use of ESF (**20**), which resulted in double derivatisation at the α -position. Subsequent hydrolysis of the boronic ester intermediate, and elimination of the chloroethane sulfonamide, will result in the respective boronic ester and ethenesulfonamide dual-warhead of **121** and **122**.

To conclude, **Figure 7.19** features all the previously used and newly proposed electrophiles and functional groups that could be included on the 4-position of driving scaffold **85**. This will result in the library of compounds labelled **123**, able to reversibly interact with, and potentially covalently modify the mutant Ser797 residue. This example study highlights the versatility of utilising these reactive fragments. Once in hand, they may be incorporated in various other scaffolds at will to generate large libraries in an efficient and divergent manner.

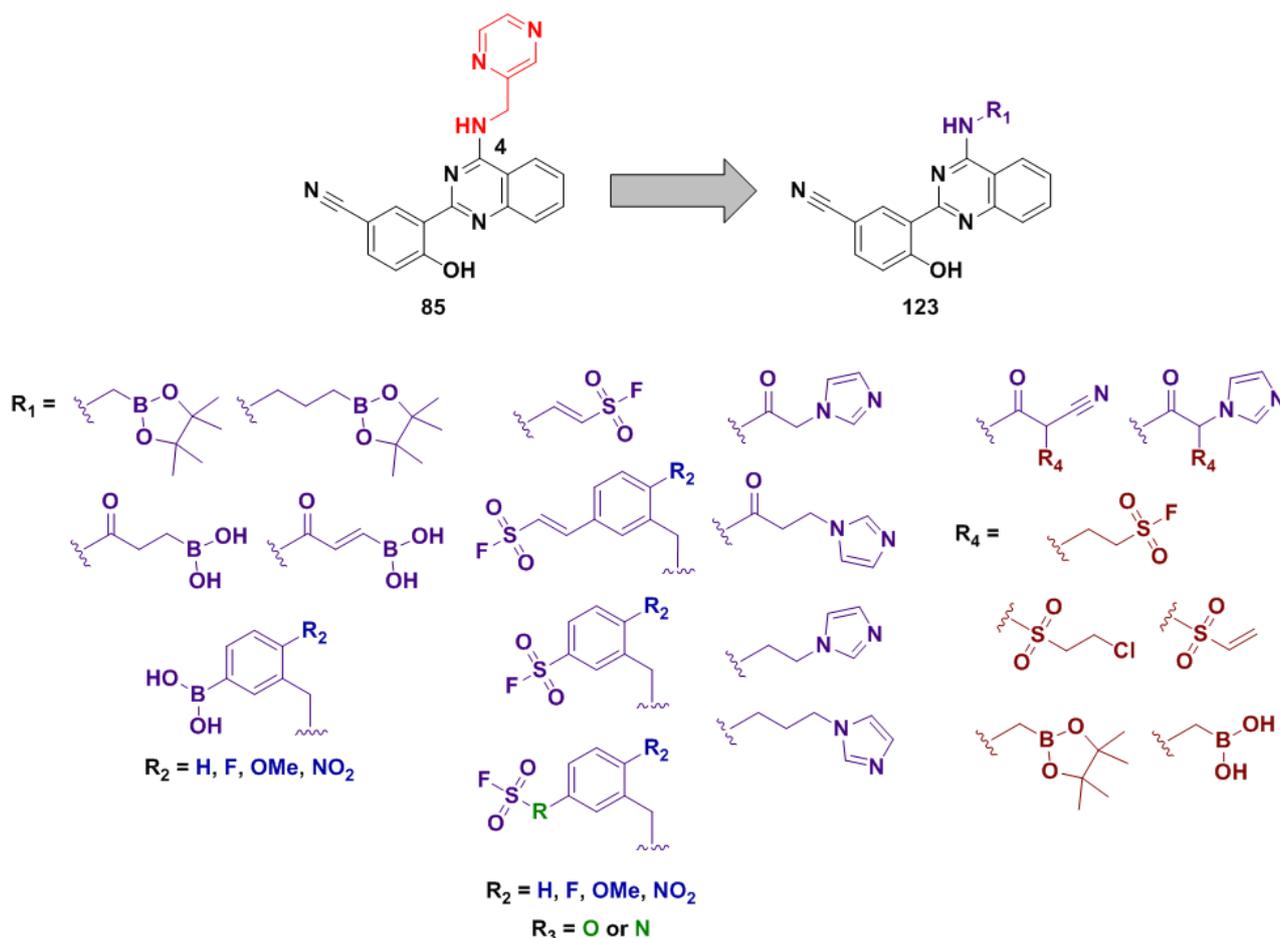
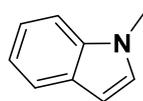


Figure 7.19: Envisaged library of inhibitors **123**, based on driving group scaffold **85** and incorporating previously used and newly proposed electrophiles and functional groups.

7.8 Supplementary Information

1-Methyl-1H-indole (**6**)²⁴



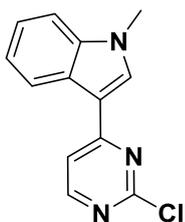
A flame-dried, 2-neck round-bottomed flask was charged with sodium hydride (60% dispersion in mineral oil, 4.10 g, 102 mmol, 2 equiv.), purged with Ar and to this was added THF (35 mL). The resulting suspension was cooled to $-10\text{ }^{\circ}\text{C}$ (acetone/ice) and a solution of indole (6.00 g, 51.2 mmol, 1 equiv.) in THF (35 mL) was added dropwise over 25 min. The solution was then allowed to warm to rt and stirred for 1 h. After recooling the solution to $-10\text{ }^{\circ}\text{C}$ (acetone/ice), a solution of MeI (8.72 g, 61.4 mmol, 1.2 equiv.) in THF (15 mL) was added dropwise over 10 min. The reaction mixture was then allowed to warm to rt and stirred for 3 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then cool to $-10\text{ }^{\circ}\text{C}$ and quenched with a saturated solution of NH_4Cl (50 mL), diluted with EtOAc (40 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc ($3 \times 30\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 20% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **6** as a clear oil (6.58 g, 50.2 mmol, 98%).

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.²⁴

Rf: 0.75 (20 % EtOAc in PE); ^1H NMR (300 MHz, CDCl_3) δ 3.83 (s, 3H, ArNCH_3), 6.57 (dd, $J = 3.1, 0.8$ Hz, 1H, ArH), 6.57 (dd, $J = 3.1, 0.8$ Hz, 1H, ArH), 7.11 (d, $J = 3.1$ Hz, 1H, ArH), 7.21 (dd, $J = 8.0, 1.0$ Hz, 1H, ArH), 7.33 (dd, $J = 6.9, 1.1$ Hz, 1H, ArH), 7.40 (dd, $J = 8.0, 1.0$ Hz, 1H, ArH), 7.75 – 7.70 (m, 1H, ArH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 32.9, 101.0, 109.3, 119.4, 121.0, 121.6, 128.9 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_9\text{H}_{10}\text{N}$: 132.0813; found: 132.0813.

3-(2-Chloropyrimidin-4-yl)-1-methyl-1H-indole (8)²³



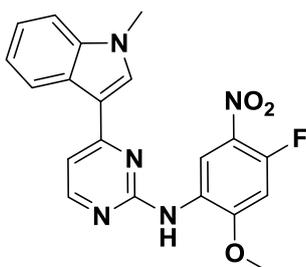
A flame-dried, 2-neck round-bottomed flask was charged with 2,4-dichloropyrimidine (8.19 g, 55.0 mmol, 1.1 equiv.), purged with Ar and dissolved in DME (40 mL). The resulting solution was cooled to 0 °C and AlCl_3 (6.67 g, 50.0 mmol, 1 equiv.) was added portionwise over 5 min.

The solution was then allowed to warm to rt and stirred for 30 min. After recooling the solution to 0 °C, a solution of 1-methyl-1H-indole (6.60 g, 50.0 mmol, 1 equiv.) in DME (35 mL) was added dropwise over 10 min, during which time the reaction mixture developed into a white-pink slurry. The reaction mixture was then incrementally increased in temperature from rt to 60 °C over 30 min and then stirred for 12 h or until complete consumption of the starting material as indicated by TLC. After allowing to cool to rt, the reaction mixture was slowly poured into vigorously stirring H_2O (80 mL) and allowed to stir for 1 h. The precipitate that formed was then collected by filtration and washed with aliquots of H_2O (3 × 40 mL) and MeOH (3 × 5 mL). Purification of the crude collected precipitate was achieved using flash column chromatography with elution gradient of 0 – 2% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **8** as a white solid (8.29 g, 34.0 mmol, 68%).

^1H NMR and MS data collected for this compound compared well with the reported literature values.²³

Rf: 0.53 (DCM); ^1H NMR (300 MHz, $\text{DMSO}-d_6$) δ 3.89 (s, 3H, ArNCH_3), 7.32 – 7.26 (m, 2H, 2 × ArH), 7.57 (dd, $J = 6.5, 2.0$ Hz, 1H, ArH), 7.81 (d, $J = 5.5$ Hz, 1H, ArH), 8.41 (dd, $J = 6.5, 2.0$ Hz, 1H, ArH), 8.48 (s, 1H, ArH), 8.52 (d, $J = 5.5$ Hz, 1H, ArH) ppm; ^{13}C NMR (151 MHz, $\text{DMSO}-d_6$) δ 33.2, 110.7, 110.8, 114.3, 121.6, 121.7, 122.7, 125.2, 134.7, 137.8, 158.7, 160.3, 164.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{12}\text{H}_9\text{N}_3\text{Cl}$: 244.0642; found: 244.0640.

N-(4-Fluoro-2-methoxy-5-nitrophenyl)-4-(1-methyl-1H-indol-3-yl)pyrimidin-2-amine (10)²³



A 2-neck round-bottomed flask was charged with 3-(2-chloropyrimidin-4-yl)-1-methyl-1H-indole (6.41 g, 26.3 mmol, 1 equiv.) and 4-fluoro-2-methoxy-5-nitroaniline (4.90 g, 26.3 mmol, 1 equiv.) and purged with Ar. To this was added 2-pentanol (80 mL) and the resulting suspension was placed at 105 °C. *p*-Toluenesulfonic acid monohydrate (1.96 g, 31.6 mmol, 1.2 equiv.) was added as a single portion and the reaction mixture was then allowed to stir for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then allowed to cool and stand for 8 h which resulted in the formation of a yellow precipitate.

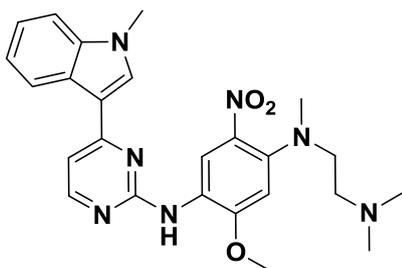
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

The precipitate was collected by filtration, washed with aliquots of acetonitrile (3×30 mL) and dried *in vacuo*. The solid was dissolved in 10% MeOH in DCM and washed with a saturated solution of NaHCO_3 (3×50 mL), a saturated solution of brine (50 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the solid was then triturated with acetonitrile and dried under vacuum to afford compound **10** as a yellow solid (9.63 g, 24.5 mmol, 93%) with no further purification required.

^1H NMR and MS data collected for this compound compared well with the reported literature values.²³

Rf: 0.74 (1% MeOH in DCM); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 3.89 (s, 3H, ArNCH_3), 3.98 (s, 3H, ArOCH_3), 7.11 (dd [app. t], $J = 7.5$ Hz, 1H, ArH), 7.29 (dd [app. t], $J = 7.1$ Hz, 1H, ArH), 7.45 – 7.41 (m, 1H, ArH), 7.47 (d, $J = 8.2$ Hz, 1H, ArH), 7.56 (d, $J = 8.2$ Hz, 1H, ArH), 8.14 (d, $J = 7.7$ Hz, 1H, ArH), 8.34 (d, $J = 6.5$ Hz, 1H, ArH), 8.67 (s, 1H, ArH), 8.75 (d, $J = 8.2$ Hz, 1H, ArH), 9.99 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 33.6, 57.6, 101.9, 107.0, 110.8, 111.1, 111.6, 114.3, 121.7, 121.9, 122.2, 122.7, 123.3, 125.2, 134.7, 138.1, 154.0, 158.7, 164.5, 165.9 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{20}\text{H}_{17}\text{N}_5\text{O}_3\text{F}$: 394.1315; found: 394.1309.

***N*1-[2-(Dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]-2-nitrobenzene-1,4-diamine (**11**)**²³



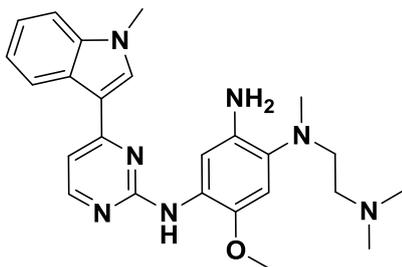
A 2-neck round-bottom flask was charged with *N*-(4-fluoro-2-methoxy-5-nitrophenyl)-4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-amine (9.13 g, 23.2 mmol, 1 equiv.), purged with Ar and dissolved in DMA (60 mL). To this solution was added DIPEA (4.24 mL, 24.4 mmol, 1.05 equiv.), followed by the slow addition of a solution of N^1,N^1,N^2 -trimethylethane-1,2-diamine (2.96 g, 29.0 mmol, 1.25 equiv.) in DMA (20 mL) at rt. The reaction

was placed at 60 °C and allowed to stir for 4 h, during which time the reaction mixture developed a deep red colour. After allowing to cool to rt, the reaction mixture was then diluted with H_2O (80 mL) and EtOAc (100 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×40 mL) and the combined organic layers were then washed with successive aliquots of a saturated solution of brine (3×40 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **11** as a red-orange powder (10.6 g, 22.3 mmol, 96%).

^1H NMR and MS data collected for this compound compared well with the reported literature values.²³

Rf: 0.43 (10% MeOH in DCM); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 2.22 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.56 (t, $J = 6.8$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.85 (s, 3H, ArNCH_3), 3.29 (t, $J = 6.8$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.86 (s, 3H, ArNCH_3), 3.95 (s, 3H, ArOCH_3), 6.85 (s, 1H, ArH), 7.12 (dd [app. t], $J = 7.6$ Hz, 1H, ArH), 7.21 (d, $J = 5.6$ Hz, 1H, ArH), 7.26 (dd [app. t], $J = 7.6$ Hz, 1H, ArH), 7.50 (d, $J = 8.0$ Hz, 1H, ArH), 8.08 (s, 1H, ArH), 8.31 (s, 1H, ArH), 8.32 (d, $J = 5.6$ Hz, 1H, ArH), 8.36 (d, $J = 8.0$ Hz, 1H, ArH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 33.0, 40.7, 45.2, 52.5, 56.1, 56.4, 102.2, 107.4, 110.4, 112.3, 119.1, 120.9, 121.5, 122.0, 122.2, 125.4, 132.3, 133.0, 137.6, 143.6, 154.7, 157.1, 160.0, 162.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{25}\text{H}_{30}\text{N}_7\text{O}_3$: 476.2410; found: 476.2403.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

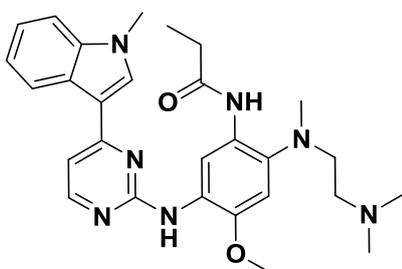
N1-[2-(Dimethylamino)ethyl]-5-methoxy-N1-methyl-N4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (12)²³

A 2-neck round-bottom flask was charged with N1-[2-(dimethylamino)ethyl]-5-methoxy-N1-methyl-N4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]-2-nitrobenzene-1,4-diamine (4.00 g, 8.41 mmol, 1 equiv.), iron (2.82 g, 50.5 mmol, 6 equiv.) and NH₄Cl (450 mg, 8.41 mmol, 1 equiv.) and purged with Ar. To this was added EtOH (75 mL) and H₂O (25 mL) and the resulting suspension was placed at 85 °C and allowed to stir for 4 h, during which time

the reaction mixture changed from brown red to a purple green colour. After allowing to cool to rt, the reaction mixture was filtered through a plug of celite, washed with EtOH (50 mL) and the solvent was removed *in vacuo*. The residue was then diluted with DCM (50 mL) and H₂O (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 30 mL) and the combined organic layers were then washed with a saturated solution of brine (50 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 15% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **12** as a brown foam (3.19 g, 7.15 mmol, 85%).

¹H NMR and MS data collected for this compound compared well with the reported literature values.²³

Rf: 0.43 (10% MeOH in DCM); IR (ATR) cm⁻¹: 3419 (m, N-H stretch); 2933 (m, C-H stretch), 2778 (m, C-H stretch), 1573 (m, N-H bend), 1444 (m, C-H bend), 1402 (m, C-H bend), 1363 (s, C-N stretch), 1097 (s, C-O stretch), 807 (s, C-H bend), 741 (s, N-H wag); ¹H NMR (300 MHz, DMSO-*d*₆) δ 2.25 (s, 6H, CH₂N(CH₃)₂), 2.46 (t, *J* = 6.6 Hz, 2H, NCH₂CH₂N), 2.62 (s, 3H, ArNCH₃), 2.93 (t, *J* = 6.6 Hz, 2H, NCH₂CH₂N), 3.75 (s, 3H, ArNCH₃), 3.87 (s, 3H, ArOCH₃), 4.92 (br s, 2H, ArNH₂), 7.16 – 7.13 (m, 1H, ArH), 7.18 (dd, *J* = 8.0, 1.0 Hz, 1H, ArH), 7.28 – 7.22 (m, 1H, ArH), 7.51 (d, *J* = 8.2 Hz, 1H, ArH), 7.53 (s, 1H, ArH), 7.79 (s, 1H, ArH), 8.27 (d, *J* = 5.4 Hz, 1H, ArH), 8.30 (s, 1H, ArH), 8.43 (d, *J* = 8.0 Hz, 1H, ArH) ppm; ¹³C NMR (75 MHz, DMSO-*d*₆) δ 33.0, 41.6, 45.3, 53.5, 56.5, 57.0, 105.3, 106.8, 109.0, 110.3, 112.4, 120.9, 122.1, 125.4, 125.5, 132.8, 133.2, 136.8, 137.6, 141.7, 157.0, 160.4, 162.1 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₅H₃₂N₇O: 446.2668; found: 446.2673.

N-(2-{[2-(Dimethylamino)ethyl](methylamino)-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl}propionamide (13)

A 2-neck round-bottom flask was charged with N1-[2-(dimethylamino)ethyl]-5-methoxy-N1-methyl-N4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (60.0 mg, 0.135 mmol, 1 equiv.) and potassium carbonate (28.0 mg, 0.203 mmol, 1.5 equiv.) and purged with Ar. To this was added acetone (4 mL) and the solution was cooled to –10 °C (acetone/ice). A solution of propionyl chloride (16.0 mg, 0.176 mmol, 1.3 equiv.) in acetone

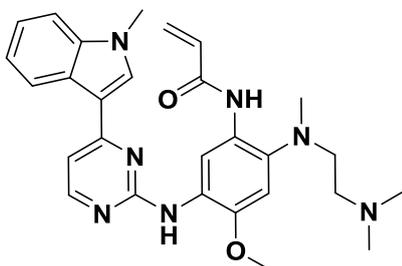
(2 mL) was then added dropwise to the solution over 5 min. The reaction mixture was then allowed to warm to rt and stirred for 2 h. The reaction mixture was then quenched and diluted with H₂O (20 mL), a saturated solution of NaHCO₃ (20 mL) and DCM (30 mL) and the organic layer was separated.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

The aqueous layer was extracted with aliquots of DCM (3 × 20 mL) and the combined organic layers were then washed with a saturated solution of brine (50 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **13** as a white foam (60 mg, 0.120 mmol, 90%).

Rf: 0.33 (10% MeOH in DCM); IR (ATR) cm⁻¹: 3419 (m, N-H stretch), 3116 (m, N-H stretch), 2953 (m, C-H stretch), 2819 (m, C-H stretch), 1670 (s, C=O stretch), 1575 (m, N-H bend), 1447 (m, C-H bend), 1402 (m, C-H bend), 1336 (s, C-N stretch), 1099 (s, C-O stretch), 805 (s, C-H bend), 739 (s, N-H wag); ¹H NMR (300 MHz, CDCl₃) δ 1.33 (t, *J* = 7.6 Hz, 3H, COCH₂CH₃), 2.29 (s, 6H, CH₂N(CH₃)₂), 2.33 (t, *J* = 5.8 Hz, 2H, NCH₂CH₂N), 2.49 (q, *J* = 7.6 Hz, 2H, COCH₂CH₃), 2.70 (s, 3H, ArNCH₃), 2.93 (t, *J* = 5.8 Hz, 2H, NCH₂CH₂N), 3.90 (s, 3H, ArNCH₃), 4.03 (s, 3H, ArNOCH₃), 6.80 (s, 1H, ArH), 7.22 (d, *J* = 5.3 Hz, 1H, ArH), 7.32 – 7.27 (m, 2H, 2 × ArH), 7.44 – 7.39 (m, 1H, ArH), 8.14 – 8.07 (m, 1H, ArH), 7.76 (s, 1H, ArH), 8.41 (d, *J* = 5.3 Hz, 1H, ArH), 9.11 (s, 1H, ArH), 9.77 (s, 2H, ArH and ArNH) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 10.1, 30.7, 33.1, 44.0, 45.5, 56.2, 56.2, 57.4, 104.6, 107.9, 109.7, 110.1, 113.7, 120.3, 120.9, 121.8, 126.0, 127.7, 129.8, 134.1, 135.2, 138.3, 157.9, 159.6, 143.9, 162.2, 171.4 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₂₈H₃₆N₇O₂: 502.2930; found: 502.2938.

N-(2-[[2-(Dimethylamino)ethyl](methylamino)-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl]acrylamide (14)²³



A 2-neck round-bottom flask was charged with N1-[2-(dimethylamino)ethyl]-5-methoxy-N1-methyl-N4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (300 mg, 0.673 mmol, 1 equiv.) and potassium carbonate (139 mg, 1.01 mmol, 1.5 equiv.) and purged with Ar. To this was added acetone (12 mL) and the solution was cooled to –40 °C (acetonitrile/CO₂). A solution of acryloyl chloride (67.0 mg, 0.740 mmol, 1.1

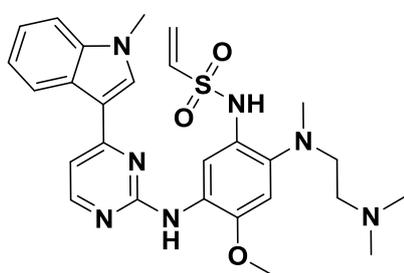
equiv.) in acetone (4 mL) was then added dropwise to the solution over 5 min. The reaction mixture was then allowed to warm to –20 °C and stirred at this temperature for 2 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H₂O (20 mL), a saturated solution of NaHCO₃ (20 mL) and DCM (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 20 mL) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **14** as a cream foam (296 mg, 0.592 mmol, 88%).

¹H NMR and MS data collected for this compound compared well with the reported literature values.²³

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Rf: 0.44 (10% MeOH in DCM); ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 2.25 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.36 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.71 (s, 3H, ArNCH_3), 2.91 (t, $J = 5.4$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.86 (s, 3H, ArNCH_3), 3.91 (s, 3H, ArOCH_3), 5.77 (dd, $J = 10.1, 1.9$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.28 (dd, $J = 16.9, 1.9$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 6.49 (dd, $J = 16.9, 10.1$ Hz, 1H, $\text{CCH}=\text{CH}_2$), 7.03 (s, 1H, ArH), 7.18 – 7.13 (m, 1H, ArH), 7.26 – 7.22 (m, 2H, $2 \times \text{ArH}$), 7.52 (d, $J = 8.0$ Hz, 1H, ArH), 7.91 (s, 1H, ArH), 8.25 (d, $J = 8.0$ Hz, 1H, ArH), 8.33 (d, $J = 5.3$ Hz, 1H, ArH), 8.68 (s, 1H, ArH), 9.16 (s, 1H, ArH), 10.18 (s, 1H, ArNH) ppm; ^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 32.9, 42.7, 44.9, 55.3, 56.0, 56.6, 105.2, 107.1, 110.5, 112.4, 113.6, 120.9, 121.3, 121.9, 125.3, 125.4, 126.1, 127.5, 132.4, 133.8, 137.5, 137.7, 146.0, 157.7, 159.9, 161.6, 162.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{34}\text{N}_7\text{O}_2$: 500.2774; found: 500.2775.

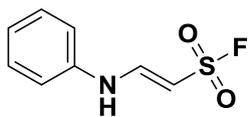
N-(2-([2-(Dimethylamino)ethyl](methyl)amino)-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)ethanesulfonamide (17)



A flame-dried, 2-neck round-bottom flask was charged with *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (150 mg, 0.337 mmol, 1 equiv.), purged with Ar and dissolved in DCM (8 mL). To this was added DIPEA (0.180 mL, 1.01 mmol, 3 equiv.) and the solution was cooled to -40 °C (acetonitrile/ CO_2). A solution of 2-bromoethane-1-sulfonyl chloride (91.0 mg, 0.438 mmol, 1.3 equiv.) in DCM (4 mL) was then added dropwise to the solution over 5 min. The reaction mixture was then allowed to warm to rt and stirred at this temperature for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (20 mL), a saturated solution of NaHCO_3 (20 mL) and DCM (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **17** as a tan solid (123 mg, 0.23 mmol, 68%).

Rf: 0.64 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3427 (m, N-H stretch), 3023 (w, C=C-H stretch), 2953 (m, C-H stretch), 2819 (m, C-H stretch), 1575 (m, N-H bend), 1447 (m, C-H bend), 1404 (m, C-H bend), 1322 (s, C-N stretch), 1147 (s, S=O stretch), 885 (s, C=C bend), 803 (s, C-H bend), 735 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 2.27 (t, $J = 5.6$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.35 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.73 (s, 3H, ArNCH_3), 2.87 (t, $J = 5.6$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.90 (s, 3H, ArNCH_3), 3.97 (s, 3H, ArOCH_3), 5.74 (d, $J = 10.0$ Hz, 1H, $\text{SCH}=\text{CH}_2$), 6.16 (d, $J = 16.7$ Hz, 1H, $\text{SCH}=\text{CH}_2$), 6.58 (dd, $J = 16.7, 10.0$ Hz, 1H, $\text{SCH}=\text{CH}_2$), 6.78 (s, 1H, ArH), 7.22 (d, $J = 5.3$ Hz, 1H, ArH), 7.33 – 7.27 (m, 2H, $2 \times \text{ArH}$), 7.44 – 7.39 (m, 1H, ArH), 7.72 (s, 1H, ArH), 8.13 – 8.07 (m, 1H, ArH), 8.39 (d, $J = 5.3$ Hz, 1H, ArH), 8.76 (s, 1H, ArNH), 9.02 (s, 1H, ArNHSO_2) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 33.3, 44.4, 44.4, 56.1, 56.5, 57.0, 104.8, 108.2, 110.3, 113.6, 113.7, 120.4, 121.1, 122.0, 125.3, 125.9, 128.1, 128.5, 134.7, 136.6, 137.2, 138.3, 145.7, 158.0, 159.6, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{34}\text{N}_7\text{O}_3\text{S}$: 536.2444; found: 536.2450.

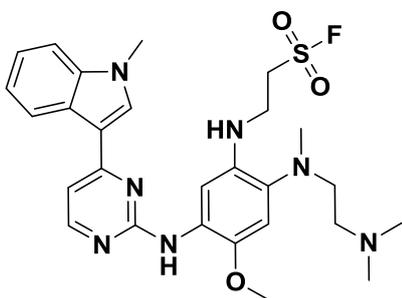
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

2-(Phenylamino)ethene-1-sulfonyl fluoride (19)²⁵

A 2-neck round-bottomed flask was charged with 2-(phenylamino)ethane-1-sulfonyl fluoride (300 mg, 1.48 mmol, 1 equiv.) and purged with Ar. To this was added CHCl_3 (10 mL), followed by the addition of MnO_2 (1.92 g, 22.1 mmol, 15 equiv.) in one portion.

The reaction mixture was allowed to stir for 5 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was filtered through a plug of celite, washed with CHCl_3 (40 mL) and the solvent was removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 25% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **19** as a white solid (205 mg, 1.02 mmol, 69%).

Rf: 0.50 (30% PE in EtOAc); ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 5.72 (d, $J = 10.8$ Hz, 1H, $\text{CH}=\text{CHS}$), 7.10 (dd [app. t], $J = 6.2$ Hz, 1H, ArH), 7.21 (d, $J = 7.8$ Hz, 2H, ArH), 7.35 (dd [app. t], $J = 7.8$ Hz, 2H, ArH), 8.10 – 7.98 (m, 1H, $\text{NHCH}=\text{CH}$), 10.48 (s, 1H, ArNH) ppm; ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 87.0, 87.2, 116.3, 123.7, 129.7, 139.6, 148.0 ppm; ^{19}F NMR (282 MHz, $\text{DMSO}-d_6$) δ 72.12 (d, $J = 2.2$ Hz) ppm; HRMS-TOF MS ESI–: m/z $[\text{M}-\text{H}]^+$ calculated for $\text{C}_8\text{H}_7\text{NO}_2\text{SF}$: 200.0182; found: 200.0191.

2-[(2-{[2-(Dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]ethane-1-sulfonyl fluoride (21)

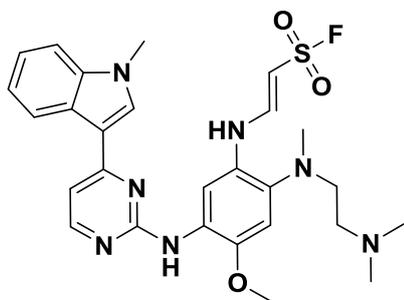
A flame-dried, 2-neck round-bottomed flask was charged with *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (100 mg, 0.224 mmol, 1 equiv.) and purged with Ar. To this was added DMF (4 mL) and the resulting solution was then cooled to -10 °C (acetone/ice), followed by the dropwise addition of a solution of ethenesulfonyl fluoride (26.0 mg, 0.235 mmol, 1.05 equiv.) in DMF (1 mL). The reaction mixture was then allowed to warm to rt and

stirred for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (30 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (3×30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5% – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **21** as a yellow solid (91.0 mg, 0.164 mmol, 73%).

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Rf: 0.40 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3418 (m, N-H stretch), 3283 (m, N-H stretch), 2920 (m, C-H stretch), 2819 (m, C-H stretch), 1579 (m, N-H bend), 1447 (m, C-H bend), 1402 (m, C-H bend), 1367 (s, C-N stretch), 1192 (s, S=O stretch), 1099 (s, C-O stretch), 1033 (s, S=O stretch), 801 (s, C-H bend), 737 (s, N-H wag); ^1H NMR (600 MHz, $\text{DMSO-}d_6$) δ 2.57 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.59 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 3.03 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.13 (s, 3H, ArNCH_3), 3.65 (br s, 2H, $\text{NHCH}_2\text{CH}_2\text{S}$), 3.79 (s, 3H ArNCH_3), 3.87 (s, 3H, ArOCH_3), 4.19 (br s, 2H, $\text{NHCH}_2\text{CH}_2\text{S}$), 5.84 (s, 1H, ArH), 6.89 (s, 1H, ArH), 7.15 (d, $J = 7.4$ Hz, 1H, ArH), 7.17 (d, $J = 5.2$ Hz, 1H, ArH), 7.25 (t, $J = 7.4$ Hz, 1H, ArH), 7.54 – 7.47 (m, 2H, $2 \times \text{ArH}$), 7.90 (s, 1H, ArH), 8.26 (d, $J = 5.2$ Hz, 1H, ArH), 8.28 (s, 1H, ArNH), 8.42 (d, $J = 7.4$ Hz, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 33.0, 38.0, 41.8, 43.0, 49.3, 49.5, 50.9, 54.4, 54.9, 56.5, 105.3, 106.4, 107.1, 110.4, 112.4, 120.8, 122.2, 125.5, 126.2, 132.7, 133.9, 136.0, 137.6, 142.0, 156.9, 160.4, 162.2 ppm; ^{19}F NMR (282 MHz, $\text{DMSO-}d_6$) δ 57.69 (t, $J = 6.1$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{35}\text{N}_7\text{O}_3\text{SF}$: 556.2506; found: 556.2494.

(E)-2-[(2-{[2-(Dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]ethene-1-sulfonyl fluoride (22)

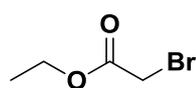


A 2-neck round-bottomed flask was charged with 2-[(2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]ethane-1-sulfonyl fluoride (70.0 mg, 0.126 mmol, 1 equiv.) and purged with Ar. To this was added CHCl_3 (5 mL), followed by the addition of MnO_2 (164 mg, 1.890 mmol, 15 equiv.) in one portion. The reaction mixture was then allowed to stir for 8 h or until complete consumption of the starting material as indicated by TLC.

The reaction mixture was filtered through a plug of celite, washed with CHCl_3 (40 mL) and the solvent was removed *in vacuo*. Purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **22** as a tan solid (52.0 mg, 94.5 μmol , 75%).

Rf: 0.73 (15 % MeOH in DCM); IR (ATR) cm^{-1} : 3414 (m, N-H stretch), 2959 (m, C-H stretch), 2850 (m, C-H stretch), 1631 (s, C=C stretch), 1577 (m, N-H bend), 1447 (m, C-H bend), 1404 (m, C-H bend), 1371 (s, C-N stretch), 1151 (s, S=O stretch), 1099 (s, C-O stretch), 1021 (s, S=O stretch), 879 (s, C=C bend), 803 (s, C-H bend), 735 (s, N-H wag); ^1H NMR (600 MHz, $\text{DMSO-}d_6$) δ 2.32 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.54 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.70 (s, 3H, ArNCH_3), 3.00 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.86 (br s, 6H, ArOCH_3 and ArNCH_3), 5.82 (d, $J = 11.2$ Hz, 1H, $\text{NHCH}=\text{CHS}$), 7.01 (d, $J = 7.4$ Hz, 1H, ArH), 7.12 (s, 1H, ArH), 7.26 – 7.19 (m, 2H, $2 \times \text{ArH}$), 7.51 (d, $J = 8.1$ Hz, 1H, ArH), 7.92 (d, $J = 11.2$ Hz, 1H, $\text{NHCH}=\text{CHS}$), 8.02 (s, 1H, ArH), 8.06 (s, 1H, ArH), 8.36 – 8.26 (m, 2H, $2 \times \text{ArH}$), 8.40 (s, 1H, ArNH), 10.20 (s, 1H, ArNH), ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 33.0, 42.0, 42.6, 44.4, 56.1, 105.1, 106.0, 107.5, 109.8, 110.4, 112.3, 120.9, 121.9, 122.2, 125.5, 126.3, 128.3, 132.9, 137.2, 137.6, 147.2, 148.6, 157.0, 160.1, 162.2 ppm; ^{19}F NMR (282 MHz, $\text{DMSO-}d_6$) δ 73.25 (d, $J = 101.8$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{33}\text{N}_7\text{O}_3\text{SF}$: 554.2350; found: 554.2352.

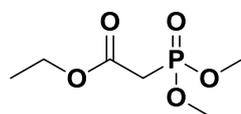
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Ethyl 2-bromoacetate (30)²⁶

A 2-neck round-bottomed flask was charged with 2-bromoacetic acid (13.9 g, 100 mmol, 1 equiv.), purged with Ar and dissolved in EtOH (100 mL). To this was added H₂SO₄ (20 mL) and the resulting solution was then allowed to stir at 85 °C for 24 h. The reaction mixture was then diluted with H₂O (100 mL) and the organic layer was separated. The organic layer was then washed with H₂O (2 × 20 mL), an ice-cold solution of 5% NaHCO₃ (20 mL) and a saturated solution of brine (20 mL) to afford compound **30** as a clear oil (14.1 g, 86.2 mmol, 86%).

¹H and ¹³C NMR data collected for this compound compared well with the reported literature values.⁷¹

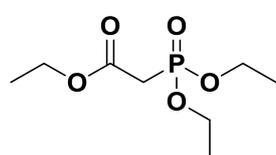
¹H NMR (400 MHz, CDCl₃) δ 1.29 (t, *J* = 7.2 Hz, 3H, CH₃CH₂O), 3.81 (s, 2H, COCH₂Br), 4.22 (q, *J* = 7.2 Hz, 2H, CH₃CH₂O) ppm; ¹³C NMR (101 MHz, CDCl₃) δ 14.1, 26.1, 62.4, 167.3 ppm; no ionization was found in MS.

Ethyl 2-(dimethoxyphosphoryl)acetate (31)²⁷

A 2-neck round-bottomed flask was charged with ethyl 2-bromoacetate (5.39 g, 32.3 mmol, 1 equiv.), purged with Ar and heated to 80 °C. Trimethyl phosphite (3.81 mL, 32.3 mmol, 1 equiv.) was added rapidly via syringe and the reaction mixture was placed at 130 °C to allow distillation of bromomethane. The reaction mixture was then allowed to stir for a further 3 h. Purification of the crude product was achieved by kugelrohr distillation. Fractions with boiling point between 145 °C – 150 °C at ~ 20 mbar were collected to afford compound **31** as a clear oil (4.72 g, 28.0 mmol, 86%).

¹H and ¹³C NMR data collected for this compound compared well with the reported literature values.⁷²

¹H NMR (300 MHz, CDCl₃) δ 1.21 (t, *J* = 7.1 Hz, 3H, CH₃CH₂O), 2.90 (d, *J* = 21.5 Hz, 2H, COCH₂P), 3.72 (s, 3H, P(OCH₃)₂), 3.74 (s, 3H, P(OCH₃)₂), 4.13 (q, *J* = 7.1 Hz, 2H, CH₃CH₂O) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 14.0, 33.4 (d, *J* = 135.1 Hz), 53.1 (*J* = 6.0 Hz), 61.6, 165.6 (d, *J* = 6.0 Hz) ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₆H₁₄O₅P: 197.0579; found: 197.0580.

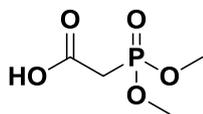
Ethyl 2-(diethoxyphosphoryl)acetate (32)²⁷

A 2-neck round-bottomed flask was charged with ethyl 2-bromoacetate (1.00 g, 5.99 mmol, 1 equiv.), purged with Ar and heated to 80 °C. Triethyl phosphite (1.03 mL, 5.99 mmol, 1 equiv.) was added rapidly via syringe and the reaction mixture was placed at 130 °C to allow distillation of bromoethane. The reaction mixture was then allowed to stir for a further 3 h. Purification of the crude product was achieved by kugelrohr distillation. Fractions with boiling point between 150 °C – 155 °C at ~ 20 mbar were collected to afford compound **32** as a clear oil (1.13 g, 5.03 mmol, 84%).

¹H and ¹³C NMR data collected for this compound compared well with the reported literature values.⁷³

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

^1H NMR (300 MHz, CDCl_3) δ 1.37 – 1.22 (m, 9H, $\text{CH}_3\text{CH}_2\text{O}$ and $\text{P}(\text{OCH}_2\text{CH}_3)_2$), 2.90 (d, $J = 1.3$ Hz, 2H, COCH_2P), 2.97 (d, $J = 1.3$ Hz, 2H, COCH_2P), 4.22 – 4.08 (m, 6H, $\text{CH}_3\text{CH}_2\text{O}$ and $\text{P}(\text{OCH}_2\text{CH}_3)_2$) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 14.1, 16.3 (d, $J = 6.2$ Hz), 34.4 (d, $J = 134.4$ Hz), 61.6, 62.8 (d, $J = 6.0$ Hz), 165.9 (d, $J = 6.0$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_8\text{H}_{18}\text{O}_5\text{P}$: 225.0892; found: 225.0892.

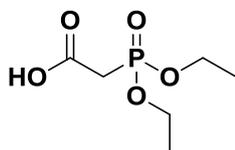
2-(Dimethoxyphosphoryl)acetic acid (33)⁷⁴

A 2-neck round-bottom flask was charged with sodium hydroxide (409 mg, 10.2 mmol, 1 equiv.) and dissolved in H_2O (5 mL). Ethyl 2-(dimethoxyphosphoryl)acetate (1.68 mL, 10.2 mmol, 1 equiv) was added dropwise to this solution and the reaction mixture was allowed to

stir for 2 h at rt. After removal of the ethanol *in vacuo*, the residue was acidified to $\text{pH} = 1$ with a solution of 1 M HCl, diluted with DCM (30 mL) and the organic layer separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, compound **33** was afforded as a white solid (1.51 g, 8.98 mmol, 88%) with no further purification required.

^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.⁷⁵

^1H NMR (300 MHz, CDCl_3) δ 2.99 (d, $J = 21.7$ Hz, 2H, COCH_2P), 3.85 – 3.75 (m, 6H, $\text{P}(\text{OCH}_3)_2$), 9.58 (br s, 1H, HOCOCH_2) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 33.3 (d, $J = 135.5$ Hz), 53.6 (d, $J = 6.4$ Hz), 167.8 (d, $J = 5.4$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_4\text{H}_{10}\text{O}_5\text{P}$: 169.0266; found: 169.0268.

2-(Diethoxyphosphoryl)acetic acid (34)⁷⁴

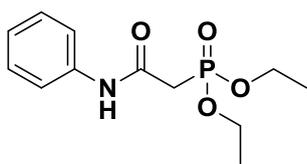
A 2-neck round-bottom flask was charged with sodium hydroxide (900 mg, 22.2 mmol, 1 equiv.) and dissolved in H_2O (11 mL). Ethyl 2-(diethoxyphosphoryl)acetate (4.40 mL, 22.2 mmol, 1 equiv.) was added dropwise to this solution and the reaction mixture was allowed to stir for 2 h at rt. After removal of the ethanol *in vacuo*, the residue was acidified to $\text{pH} =$

1 with a solution of 1 M HCl, diluted with DCM (50 mL) and the organic layer separated. The aqueous layer was extracted with aliquots of DCM (3×30 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, compound **34** was afforded as a clear oil (4.14 g, 21.1 mmol, 91%) with no further purification required.

^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.⁷⁶

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 1.23 ($2 \times t$, $J = 7.1$, 6H, $2 \times \text{POCH}_2\text{CH}_3$), 2.96 (d, $J = 21.4$ Hz, 2H, COCH_2P), 4.03 ($2 \times q$, $J = 7.1$ Hz, 4H, $2 \times \text{POCH}_2\text{CH}_3$) ppm; ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 16.1 (d, $J = 6.2$ Hz), 33.9 (d, $J = 130.7$ Hz), 61.8 (d, $J = 6.2$ Hz), 167.2 (d, $J = 5.7$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_6\text{H}_{14}\text{O}_5\text{P}$: 197.0579; found: 197.0579.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Diethyl [2-oxo-2-(phenylamino)ethyl]phosphonate (36)

A 2-neck round-bottomed flask was charged with 1,1'-carbonyldiimidazole (322 mg, 1.99 mmol, 1.85 equiv.), purged with Ar and dissolved in DCM (10 mL). The resulting solution was then cooled to 0 °C, followed by the dropwise addition of a solution of 2-(diethoxyphosphoryl)acetic acid (390 mg, 1.99 mmol, 1.85 equiv.) in DCM (4 mL).

The reaction mixture was then allowed to warm to rt and stirred for 1 h before being recooled to 0 °C. A solution of aniline (100 mg, 1.07 mmol, 1 equiv.) in DCM (1 mL) was then added dropwise to the solution and the reaction mixture was then allowed to warm to rt and stir for 12 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H₂O (20 mL), a saturated solution of NaHCO₃ (20 mL) and DCM (10 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 70 – 90% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **36** as a clear oil (272 mg, 1.00 mmol, 94%).

¹H and ¹³C NMR and MS data collected for this compound compared well with the reported literature values.⁷⁷

Rf: 0.33 (EtOAc); ¹H NMR (300 MHz, CDCl₃) δ 1.34 (2 × t, *J* = 7.1 Hz, 6H, 2 × POCH₂CH₃), 3.05 (d, *J* = 21.0 Hz, 2H, COCH₂P), 4.18 (2 × q, *J* = 7.1 Hz, 4H, 2 × POCH₂CH₃), 7.06 – 6.99 (m, 1H, ArH), 7.22 (dd [app. t], *J* = 7.9 Hz, 2H, 2 × ArH), 7.49 (m, 2H, 2 × ArH), 9.20 (s, 1H, ArNH) ppm; ¹³C NMR (101 MHz, CDCl₃) δ 16.4 (d, *J* = 6.1 Hz), 36.4 (d, *J* = 129.8 Hz), 63.1 (d, *J* = 6.6 Hz), 119.9, 124.3, 128.9, 138.0, 162.3 (d, *J* = 4.1 Hz) ppm; ³¹P NMR (121 MHz, CDCl₃) δ 23.40 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₂H₁₉NO₄P: 272.1052; found: 272.1050.

Ethyl hydrogen [2-oxo-2-(phenylamino)ethyl]phosphonate (37)²⁹

A microwave vial was charged with diethyl [2-oxo-2-(phenylamino)ethyl]phosphonate (140 mg, 0.516 mmol, 1 equiv.), sodium iodide (93.0 mg, 0.619 mmol, 1.2 equiv.), magnesium chloride (59.0 mg, 0.619 mmol, 1.2 equiv.) and purged with Ar. To this was added acetonitrile (5 mL) and after capping the vial the resulting suspension was warmed

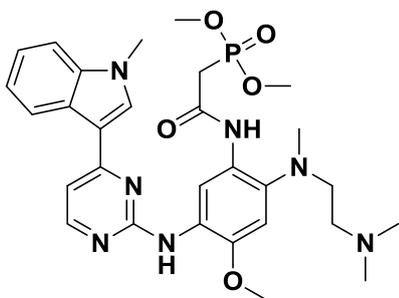
to 60 °C and allowed to stir for 12 h or until complete consumption of the starting material as indicated by TLC during which time a cream precipitate formed. The reaction mixture was then allowed to cool to rt and the white precipitate was collected by filtration. The solid was suspended in dioxane (10 mL) and a solution of 4 M HCl in dioxane (0.15 mL, 0.619 mmol, 1.2 equiv.) was then added dropwise and the reaction mixture was allowed to stir for 1 h. The reaction mixture was then filtered and after removal of the solvent *in vacuo*, compound **37** was afforded as a tan solid (95.0 mg, 0.392 mmol, 76%) with no further purification required.

Rf: 0.1 (10% MeOH in DCM); ¹H NMR (300 MHz, DMSO-*d*₆) δ 1.12 (t, *J* = 7.0 Hz, 3H, POCH₂CH₃), 3.10 (d, *J* = 20.4 Hz, 2H, COCH₂P), 3.96 – 3.81 (m, 2H, POCH₂CH₃), 7.09 (dd [app. t], *J* = 7.3 Hz, 1H, ArH), 7.31 (dd [app. t], *J* = 7.8 Hz, 2H, 2 × ArH), 7.74 (d, *J* = 7.8 Hz, 2H, 2 × ArH), 11.12 (s, 1H, ArNH) ppm;

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 16.6, 29.0, 59.8, 120.2, 124.1, 128.6, 138.3, 169.8 ppm; ^{31}P NMR (121 MHz, $\text{DMSO-}d_6$) δ 12.82 ppm; HRMS-TOF MS ESI-: m/z $[\text{M}-\text{H}]^+$ calculated for $\text{C}_{10}\text{H}_{13}\text{NO}_4\text{P}$: 242.0582; found: 252.0588.

Dimethyl {2-[(2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]-2-oxoethyl}phosphonate (39)



A 2-neck round-bottomed flask was charged with 1,1'-carbonyldiimidazole (173 mg, 1.07 mmol, 1.85 equiv.), purged with Ar and dissolved in DCM (20 mL). The resulting solution was then cooled to 0 °C, followed by the dropwise addition of a solution of 2-(dimethoxyphosphoryl)acetic acid (185 mg, 1.07 mmol, 1.85 equiv.) in DCM (5 mL). The reaction mixture was then allowed to warm to rt and stirred for 1 h before being recooled to 0 °C. A solution of N1-[2-(dimethylamino)ethyl]-5-methoxy-N1-methyl-N4-[4-(1-methyl-1H-

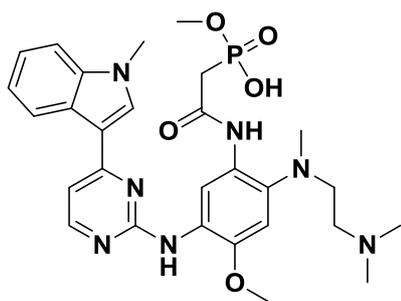
indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (257 mg, 0.579 mmol, 1 equiv.) in DCM (6 mL) was then added dropwise to the solution and the reaction mixture was then allowed to warm to rt and stir for 12 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (20 mL), a saturated solution of NaHCO_3 (20 mL) and DCM (10 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 \times 20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **39** as an off-white foam (296 mg, 0.498 mmol, 86%).

Rf: 0.38 (10 % MeOH in DCM); ^1H NMR (300 MHz, CDCl_3) δ 2.32 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.39 – 2.34 (m, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.72 (s, 3H, ArNCH_3), 2.77 (s, 2H, COCH_2P), 2.98 – 2.95 (m, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.84 (s, 3H, $\text{PO}(\text{OCH}_3)_2$), 3.87 (s, 3H, $\text{PO}(\text{OCH}_3)_2$), 3.91 (s, 3H, ArNCH_3), 4.03 (s, 3H, ArOCH_3), 6.81 (s, 1H, *ArH*), 7.22 (d, $J = 5.4$ Hz, 1H, *ArH*), 7.32 – 7.28 (m, 2H, 2 \times *ArH*), 7.43 (dd, $J = 6.0, 2.0$ Hz, 1H, *ArH*), 7.73 (s, 1H, *ArH*), 8.13 – 8.10 (m, 1H, *ArH*), 8.40 (d, $J = 5.4$ Hz, 1H, *ArH*), 8.95 (s, 1H, *ArH*), 9.66 (s, 1H, *ArNH*), 10.17 (s, 1H, *ArNH*) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 33.1 (d, $J = 10.2$ Hz), 35.6, 37.4, 44.9, 45.4, 53.3 (d, $J = 6.3$ Hz), 55.7, 56.2, 57.3, 104.6, 108.0, 110.2, 113.6, 116.6, 120.4, 121.0, 121.9, 126.0, 127.7, 130.1, 134.3, 136.3, 138.3, 144.6, 157.9, 159.6, 161.4, 162.2 ppm; ^{31}P NMR (121 MHz, CDCl_3) δ 25.53 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{29}\text{H}_{39}\text{N}_7\text{O}_5\text{P}$: 596.2750; found: 596.2742.

Methyl hydrogen {2-[(2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]-2-oxoethyl}phosphonate (40)³⁰

A 2-neck round-bottomed flask was charged with dimethyl {2-[(2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]-2-oxoethyl}phosphonate (150 mg, 0.252 mmol, 1 equiv.), purged with Ar and dissolved in *tert*-butylamine (6 mL).

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



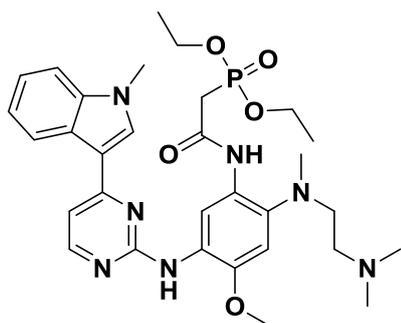
The resulting solution was then warmed to 55 °C and allowed to stir for 16 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then allowed to cool to rt, during which time a white precipitate formed. After removal of the solvent *in vacuo*, the white solid was suspended in dioxane (10 mL) and a solution of 4 M HCl in dioxane (756 μ L, 0.302 mmol, 1.2 equiv.) was then added dropwise and the reaction mixture was allowed to stir for 1 h. The reaction mixture was then filtered and after removal of the solvent *in vacuo*, compound **40** was afforded as an off-white solid (95.0 mg, 0.164 mmol, 65%) with no further purification required.

Rf: 0.11 (30 % MeOH in DCM); IR (ATR) cm^{-1} : 3363 (m, N-H stretch), 3267 (m, P-OH stretch), 2976 (m, C-H stretch), 1629 (s, C=O stretch), 1565 (m, N-H bend), 1466 (m, C-H bend), 1421 (m, C-H bend), 1328 (s, C-N stretch), 1207 (s, P=O stretch), 1097 (s, C-O stretch), 799 (s, C-H bend), 745 (s, N-H wag); ^1H NMR (600 MHz, CD_3OD) δ 2.42 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.74 – 2.68 (m, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.79 (s, 2H, COCH_2P), 3.14 (s, 3H, ArNCH_3), 3.62 – 3.57 (m, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.81 (s, 3H, POOCH_3), 3.84 (s, 3H, ArNCH_3), 3.93 (s, 3H, ArOCH_3), 7.08 (d, $J = 8.0$ Hz, 1H, ArH), 7.35 – 7.24 (m, 2H, $2 \times \text{ArH}$), 7.42 (d, $J = 8.0$ Hz, 1H, ArH), 7.59 (dd, $J = 8.0, 5.3$ Hz, 1H, ArH), 8.06 – 8.01 (m, 1H, ArH), 7.67 (s, 1H, ArH), 8.44 – 8.38 (m, 1H, ArH), 8.85 (d, $J = 5.3$ Hz, 1H, ArH), 9.13 (s, 1H, ArNH), 10.13 (s, 1H, ArNH) ppm; ^{13}C NMR (151 MHz, CD_3OD) δ 36.2, 39.4, 43.7, 46.6, 52.7, 55.7, 56.9, 58.1, 103.4, 106.3, 111.6, 112.1, 114.6, 118.9, 120.1, 122.1, 123.8, 130.1, 130.5, 134.3, 137.7, 139.3, 146.1, 158.3, 160.0, 163.5, 164.6 ppm; ^{31}P NMR (162 MHz, CD_3OD) δ 15.93 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{37}\text{N}_7\text{O}_5\text{P}$: 582.2594; found: 582.2603.

Diethyl {2-[(2-[(2-(dimethylamino)ethyl](methyl)amino)-4-methoxy-5-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino]phenyl)amino]-2-oxoethyl}phosphonate (41)

A 2-neck round-bottomed flask was charged with 1,1'-carbonyldiimidazole (81.0 mg, 0.498 mmol, 1.85 equiv.), purged with Ar and dissolved in DCM (10 mL). The resulting solution was then cooled to 0 °C, followed by the dropwise addition of a solution of 2-(diethoxyphosphoryl)acetic acid (98.0 mg, 0.498 mmol, 1.85 equiv.) in DCM (4 mL). The reaction mixture was then allowed to warm to rt and stirred for 1 h before being recooled to 0 °C. A solution of *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (120 mg, 0.269 mmol, 1 equiv.) in DCM (3 mL) was then added dropwise to the solution and the reaction mixture was then allowed to warm to rt and stir for 12 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (20 mL), a saturated solution of NaHCO_3 (20 mL) and DCM (10 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered.

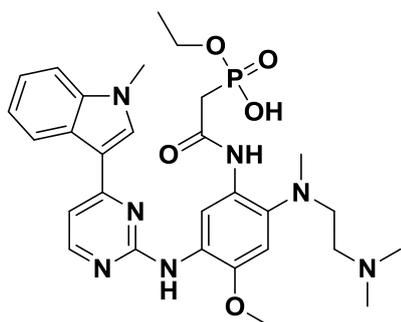
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797



After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 20% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **41** as a tan solid (147 mg, 0.237 mmol, 88%).

Rf: 0.42 (12.5% MeOH in DCM); ^1H NMR (600 MHz, CDCl_3) δ 1.26 (t, J = 6.8 Hz, 6H, $\text{PO}(\text{OCH}_2\text{CH}_3)_2$), 2.32 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.45 (t, J = 5.1 Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.62 (s, 3H, ArNCH_3), 2.94 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.13 (d, J = 21.4 Hz, 2H, COCH_2P), 3.81 (s, 3H, ArNCH_3), 3.91 (s, 3H, ArOCH_3), 4.15 – 4.09 (m, 4H, $\text{PO}(\text{OCH}_2\text{CH}_3)_2$), 6.71 (s, 1H, ArH), 7.12 (d, J = 5.1 Hz, 1H, ArH), 7.24 – 7.18 (m, 2H, $2 \times \text{ArH}$), 7.33 (d, J = 7.6 Hz, 1H, ArH), 7.66 (s, 1H, ArH), 8.05 (d, J = 7.1 Hz, 1H, ArH), 8.30 (d, J = 5.1 Hz, 1H, ArH), 8.80 (s, 1H, ArH), 9.59 (s, 1H, ArNH), 9.92 (s, 1H, ArNH) ppm; ^{13}C NMR (151 MHz, CDCl_3) δ 16.4 (d, J = 6.1 Hz), 33.0, 37.1 (d, J = 133.2 Hz), 43.9, 44.7, 54.6, 56.0, 56.5, 62.5 (d, J = 6.1 Hz), 104.2, 107.9, 110.0, 110.2, 113.4, 120.3, 120.9, 121.8, 125.8, 127.4, 128.7, 134.6, 138.1, 144.5, 157.6, 159.5, 161.7 (d, J = 5.5 Hz), 162.1 ppm; ^{31}P NMR (243 MHz, CDCl_3) δ 23.10 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{31}\text{H}_{42}\text{N}_7\text{O}_5\text{P}$: 623.2925; found: 623.2920.

Ethyl hydrogen {2-[(2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]-2-oxoethyl}phosphonate (42)²⁹



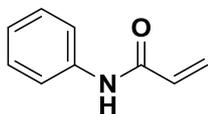
A microwave vial was charged with diethyl {2-[(2-{[2-(dimethylamino)ethyl](methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)amino]-2-oxoethyl}phosphonate (152 mg, 0.244 mmol, 1 equiv.), sodium iodide (44.0 mg, 0.292 mmol, 1.2 equiv.), magnesium chloride (27.0 mg, 0.292 mmol, 1.2 equiv.) and purged with Ar. To this was added acetonitrile (5 mL) and after capping the vial the resulting suspension was warmed to 60 °C and allowed to stir for 12 h or until complete

consumption of the starting material as indicated by TLC during which time a cream precipitate formed. The reaction mixture was then allowed to cool to rt and the white precipitate was collected by filtration. The solid was suspended in dioxane (10 mL) and a solution of 4 M HCl in dioxane (730 μL , 0.292 mmol, 1.2 equiv.) was then added dropwise and the reaction mixture was allowed to stir for 1 h. The reaction mixture was then filtered and after removal of the solvent *in vacuo*, compound **42** was afforded as a tan solid (93.0 mg, 0.156 mmol, 64%) with no further purification required.

Rf: 0.16 (30% MeOH in DCM); IR (ATR) cm^{-1} : 3375 (m, N-H stretch), 3287 (m, P-OH stretch), 2978 (m, C-H stretch), 1633 (s, C=O stretch), 1522 (m, N-H bend), 1452 (m, C-H bend), 1417 (m, C-H bend), 1371 (s, C-N stretch), 1229 (s, P=O stretch), 1089 (s, C-O stretch), 803 (s, C-H bend), 745 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 1.28 – 1.20 (m, 3H, $\text{POOCH}_2\text{CH}_3$), 2.74 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.88 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 3.13 (br s, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.27 – 3.25 (m, 2H, COCH_2P), 3.30 (s, 3H, ArNCH_3), 3.80 (s, 3H, ArNCH_3), 3.87 (s, 3H, ArOCH_3), 4.11 – 4.00 (m, 2H, $\text{POOCH}_2\text{CH}_3$), 6.56 – 6.48 (m, 1H, ArH), 6.70 (s, 1H, ArH), 6.73 (s, 1H, ArH), 6.82 – 6.75 (m, 1H, ArH), 7.04 (s, 1H, ArH), 7.13 – 7.07 (m, 1H, ArH), 7.58 (dd, J = 7.8, 1.4 Hz, 1H, ArH), 7.83 (s, 1H, ArH), 8.06 (s, 1H, ArNH), 8.44 (s, 1H, ArNH) ppm;

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

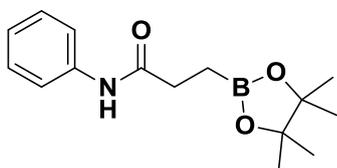
^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 17.1, 42.6, 50.0, 56.5, 60.1, 105.2, 105.7, 107.3, 110.8, 112.8, 114.7, 116.6, 121.4, 122.5, 125.9, 133.8, 133.9, 157.5, 138.1, 157.8, 158.1, 160.8, 162.7, 170.3 ppm; ^{31}P NMR (162 MHz, $\text{DMSO-}d_6$) δ 12.53 ppm HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{29}\text{H}_{39}\text{N}_7\text{O}_5\text{P}$: 596.2750; found: 596.2769.

N-Phenylacrylamide (43)

A flame-dried, 2-neck round-bottom flask was charged with aniline (1.00 g, 10.7 mmol, 1 equiv.), purged with Ar and dissolved in DCM (30 mL). DIPEA (2.24 mL, 12.9 mmol, 1.2 equiv.) was added and the solution was cooled to $-40\text{ }^\circ\text{C}$ (acetone/trile/ CO_2). A solution of acryloyl chloride (1.07 g, 11.8 mmol, 1.1 equiv.) in DCM (10 mL) was then added dropwise to the solution over 10 min. The reaction mixture was then allowed to warm to $-20\text{ }^\circ\text{C}$ and stirred at this temperature for 2 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (20 mL), a saturated solution of NaHCO_3 (20 mL) and DCM (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM ($3 \times 30\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 10 – 25% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **43** as a white solid (1.37 g, 9.34 mmol, 87%).

^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.⁷⁸

Rf: 0.24 (20% EtOAc in PE); ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 5.75 (dd, $J = 10.1, 2.1\text{ Hz}$, 1H, $\text{CCH}=\text{CH}_2$), 6.27 (dd, $J = 17.0, 2.1\text{ Hz}$, 1H, $\text{CCH}=\text{CH}_2$), 6.45 (dd, $J = 17.0, 10.1\text{ Hz}$, 1H, $\text{CCH}=\text{CH}_2$), 7.09 – 7.04 (m, 1H, ArH), 7.35 – 7.29 (m, 2H), 7.70 – 7.65 (m, 2H, $2 \times \text{ArH}$), 10.13 (s, 1H, ArNH) ppm; ^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 119.3, 123.5, 126.8, 128.8, 131.9, 139.0, 163.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_9\text{H}_{10}\text{NO}$: 148.0762; found: 148.0753.

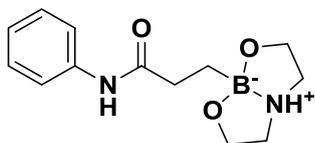
N-Phenyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanamide (44)³¹

A flame-dried Schlenk tube was charged with *N*-phenylacrylamide (100 mg, 0.679 mmol, 1 equiv.), bis(pinacolato)diboron (207 mg, 0.815 mmol, 1.2 equiv.), CuCl (7.00 mg, $67.9\text{ }\mu\text{mol}$, 10 mol%), potassium *tert*-butoxide (23.0 mg, 0.204 mmol, 30 mol%) and xantphos (38.0 mg, $67.9\text{ }\mu\text{mol}$, 10 mol%) and purged with Ar. To this was added EtOH (6 mL) and the resulting grey suspension was then allowed to stir for 6 h, filtered through a plug of celite and washed with EtOAc (30 mL). After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 10 – 25% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **44** as a clear oil (151 mg, 0.549 mmol, 81%).

Rf: 0.30 (20% EtOAc in PE); ^1H NMR (400 MHz, CDCl_3) δ 1.11 (t, $J = 7.3\text{ Hz}$, 2H $\text{COCH}_2\text{CH}_2\text{B}$), 1.24 (s, 12H, $(\text{CH}_3)_2\text{CC}(\text{CH}_3)_2$), 2.47 (t, $J = 7.3\text{ Hz}$, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 7.04 (t, $J = 7.4\text{ Hz}$, 1H, ArH), 7.30 – 7.22 (m, 2H, $2 \times \text{ArH}$), 7.52 – 7.46 (m, 2H, $2 \times \text{ArH}$), 7.83 (s, 1H, ArNH) ppm;

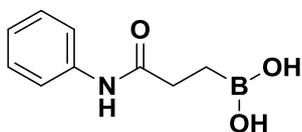
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

^{13}C NMR (101 MHz, CDCl_3) δ 24.9, 32.1, 75.0, 83.5, 119.7, 123.9, 128.9, 138.4, 172.5 ppm; ^{11}B NMR (96 MHz, CDCl_3) δ 34.49 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{15}\text{H}_{23}\text{NO}_3\text{B}$: 276.1171; found: 276.1171.

N-phenyl-3-(tetrahydro-8 λ^4 -[1,3,2]oxazaborolo[2,3-*b*][1,3,2]oxazaborol-8-yl)propanamide (45)³²

A 2-neck round-bottom flask was charged with N-phenyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanamide (360 mg, 1.31 mmol, 1 equiv.), purged with Ar and dissolved in diethyl ether (5 mL). To this was added diethanolamine (151 mg, 1.44 mmol, 1.1 equiv.) and the resulting solution was left to stir for 8 h or until complete consumption of the starting material as indicated by TLC, during which time a white precipitate formed. The precipitate was collected by filtration, washed with diethyl ether (3 \times 20 mL) and dried under vacuum to afford compound **45** as a white solid (223 mg, 0.852 mmol, 65%).

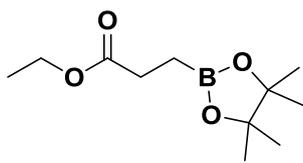
Rf: 0.63 (50% EtOAc in PE); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 0.60 – 0.49 (m, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 2.24 – 2.14 (m, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 2.80 – 2.69 (m, 2H, $\text{OCH}_2\text{CH}_2\text{NH}$), 3.09 – 2.94 (m, 2H, $\text{OCH}_2\text{CH}_2\text{NH}$), 3.66 – 3.55 (m, 2H, $\text{OCH}_2\text{CH}_2\text{NH}$), 3.78 – 3.67 (m, 2H, $\text{OCH}_2\text{CH}_2\text{NH}$), 6.81 (s, 1H, $\text{CH}(\text{NH})\text{BCH}$), 7.03 – 6.94 (m, 1H, ArH), 7.31 – 7.19 (m, 2H, 2 \times ArH), 7.62 – 7.53 (m, 2H, 2 \times ArH), 9.81 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 33.5, 50.7, 62.3, 118.9, 122.5, 128.5, 139.8, 174.9 ppm; ^{11}B NMR (96 MHz, $\text{DMSO-}d_6$) δ 13.01 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{13}\text{H}_{20}\text{N}_2\text{O}_3\text{B}$: 263.1567; found: 263.1562.

(3-Oxo-3-(phenylamino)propyl)boronic acid (46)³²

A 2-neck round-bottom flask was charged with N-phenyl-3-(tetrahydro-8 λ^4 -[1,3,2]oxazaborolo[2,3-*b*][1,3,2]oxazaborol-8-yl)propanamide (70 mg, 0.267 mmol, 1 equiv.) and dissolved in diethyl ether (3 mL). To this was added 0.1 M HCl (5 mL) and the resulting solution was left to stir for 2 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (5 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3 \times 10 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the residue was dried under vacuum to afford compound **46** as a white solid (45 mg, 0.232 mmol, 87%) with no further purification required.

Rf: 0.64 (EtOAc); ^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 0.85 (t, $J = 7.3$ Hz, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 2.35 (t, $J = 7.3$ Hz, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 6.99 (dd [app. t], $J = 7.4$ Hz, 1H, ArH), 7.30 – 7.21 (m, 2H, 2 \times ArH), 7.50 (s, 2H, $\text{B}(\text{OH})_2$), 7.58 (d, $J = 7.6$ Hz, 2H, 2 \times ArH), 9.77 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 31.6, 118.9, 122.7, 128.6, 139.5, 173.1 ppm; ^{11}B NMR (96 MHz, $\text{DMSO-}d_6$) δ 33.13 ppm; HRMS-TOF MS ESI–: m/z $[\text{M}-\text{H}]^+$ calculated for $\text{C}_9\text{H}_{11}\text{NO}_3\text{B}$: 192.0832; found: 192.0836.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

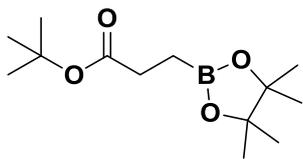
Ethyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanoate (50)³¹

A flame-dried schlenk tube was charged with ethyl acrylate (300 mg, 3.00 mmol, 1 equiv.), bis(pinacolato)diboron (913 mg, 3.60 mmol, 1.2 equiv.), CuCl (30.0 mg, 0.300 mmol, 10 mol%), potassium *tert*-butoxide (101 mg, 0.900 mmol, 30 mol%) and xantphos (174 mg, 0.300 mmol, 10 mol%) and purged with Ar. To this was added

EtOH (18 mL) and the resulting grey suspension was then allowed to stir for 6 h, filtered through a plug of celite and washed with EtOAc (30 mL). After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 1 – 2.5% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **50** as a clear oil (602 mg, 2.64 mmol, 88%).

¹H, ¹³C and ¹¹B NMR and MS data collected for this compound compared well with the reported literature values.⁷⁹

Rf: 0.37 (2.5% EtOAc in PE); ¹H NMR (300 MHz, CDCl₃) δ 0.97 (t, *J* = 7.5 Hz, 2H, COCH₂CH₂B), 1.23 – 1.17 (m, 15H, CH₃CH₂CO and (CH₃)₂CC(CH₃)₂), 2.38 (t, *J* = 7.5 Hz, 2H, COCH₂CH₂B), 4.07 (q, *J* = 7.1 Hz, 2H, CH₃CH₂CO) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 14.3, 24.6, 24.8, 28.9, 60.3, 83.3, 174.7 ppm; ¹¹B NMR (96 MHz, cdcl₃) δ 34.45 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₁H₂₂O₄B: 229.1611; found: 229.1597.

***tert*-Butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanoate (53)**³⁷

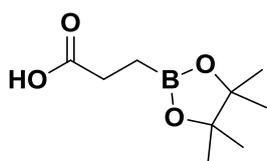
A flame-dried Schlenk tube was charged with CuCl (34.0 mg, 0.340 mmol, 5 mol%), potassium *tert*-butoxide (76.0 mg, 0.680 mmol, 10 mol%) and xantphos (200 mg, 0.340 mmol, 5 mol%), purged with Ar and to this was added THF (6 mL). The resulting suspension was then allowed to stir for 30 min at rt, followed by the dropwise addition

of a solution of bis(pinacolato)diboron (1.81 g, 7.14 mmol, 1.05 equiv.) in THF (5 mL). After stirring for a further 30 min at rt, *tert*-butyl acrylate (1.00 mL, 6.80 mmol, 1 equiv.) was added dropwise, followed by the addition of MeOH (0.5 mL) during which time a slight exotherm was witnessed. The resulting grey slurry was then allowed to stir for 5 h, filtered through a plug of celite and washed with EtOAc (30 mL). After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 2 – 10% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **53** as a clear liquid (1.71 g, 6.66 mmol, 98%).

¹H and ¹³C NMR data collected for this compound compared well with the reported literature values.³⁷

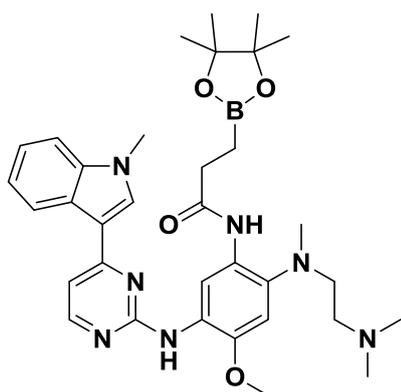
Rf: 0.38 (5% EtOAc in PE); ¹H NMR (300 MHz, CDCl₃) δ 0.92 (t, *J* = 7.5 Hz, 2H, COCH₂CH₂B), 1.20 (s, 12H, (CH₃)₂CC(CH₃)₂), 1.39 (s, 9H, (CH₃)₃CCO), 2.31 (t, *J* = 7.5 Hz, 2H, COCH₂CH₂B) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 24.9, 28.2, 30.0, 79.8, 83.2, 174.1 ppm; ¹¹B NMR (96 MHz, CDCl₃) δ 34.45 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₁₂H₂₀O₅B: 255.1404; found: 255.1400.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

3-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)propanoic acid (51)³⁶

A 2-neck round-bottomed flask was charged with *tert*-butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanoate (500 mg, 1.95 mmol, 1 equiv.), purged with Ar and dissolved in DCM (15 mL). This was followed by the dropwise addition of trifluoroacetic acid (5 mL) and the reaction mixture was then allowed to stir for 6 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, the residue was dried under vacuum to afford compound **51** as a brown oil (365 mg, 1.83 mmol, 94%) with no further purification required.

Rf: 0.06 (10% EtOAc in PE); ¹H NMR (300 MHz, CDCl₃) δ 1.01 (t, *J* = 7.3 Hz, 2H, COCH₂CH₂B), 1.23 (s, 12H, (CH₃)₂CC(CH₃)₂), 2.50 (t, *J* = 7.3 Hz, 2H, COCH₂CH₂B), 10.97 (br s, 1H, HOOCCO) ppm; ¹³C NMR (75 MHz, CDCl₃) δ 24.6, 28.6, 53.5, 84.1, 181.5 ppm; ¹¹B NMR (96 MHz, CDCl₃) δ 34.71 ppm; HRMS-TOF MS ESI+: *m/z* [M+H]⁺ calculated for C₉H₁₈O₄B: 200.1334; found: 200.1338.

N-(2-[[2-(Dimethylamino)ethyl](methyl)amino]-4-methoxy-5-[[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino]phenyl)-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanamide (47)

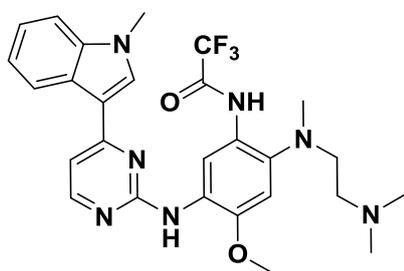
A flame dried 2-neck round-bottom flask was charged with 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)propanoic acid (45.0 mg, 0.224 mmol, 2 equiv.), purged with Ar and dissolved in DCM (5 mL). The resulting solution was cooled to 0 °C and this was followed by the portion wise addition of 1-hydroxybenzotriazole monohydrate (34.0 mg, 0.224 mmol, 2 equiv.) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (43.0 mg, 0.224 mmol, 2 equiv.) successively over 5 min. After allowing to stir for 30 min at 0 °C, a solution of *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (50.0 mg, 0.112 mmol, 1 equiv.) in DCM (3 mL) was added dropwise. The reaction mixture was stirred at 0 °C for 1 h and then allowed to warm to rt and stirred for a further 8 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H₂O (30 mL), a saturated solution of NaHCO₃ (30 mL) and DCM (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3 × 30 mL) and the combined organic layers were then washed with H₂O (30 mL), a solution of 1 M NaOH (2 × 20 mL), a saturated solution of brine (40 mL), dried over MgSO₄ and filtered. After removal of the solvent *in vacuo*, the residue was dried under vacuum to afford compound **47** as a white foam (51.0 mg, 80.8 μmol, 72%) with no further purification tolerated.

Rf: 0.53 (10% MeOH in DCM); IR (ATR) cm⁻¹: 3421 (m, N-H stretch), 3122 (m, N-H stretch), 2961 (m, C-H stretch), 2852 (m, C-H stretch), 1672 (s, C=O stretch), 1577 (m, N-H bend), 1450 (m, C-H bend), 1402 (m, C-H bend), 1342 (s, C-N stretch), 1099 (s, C-O stretch), 805 (s, C-H bend), 739 (s, N-H wag);

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

^1H NMR (300 MHz, CDCl_3) δ 1.21 (s, 12H, $(\text{CH}_3)_2\text{CC}(\text{CH}_3)_2$), 1.27 (t, $J = 6.2$ Hz, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 2.30 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.35 (t, $J = 5.9$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.57 (t, $J = 6.2$ Hz, 2H, $\text{COCH}_2\text{CH}_2\text{B}$), 2.68 (s, 3H, ArNCH_3), 2.93 (t, $J = 5.9$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.87 (s, 3H, ArNCH_3), 4.03 (s, 3H, ArOCH_3), 6.78 (s, 1H, ArH), 7.20 (d, $J = 5.3$ Hz, 1H, ArH), 7.31 – 7.23 (m, 2H, $2 \times \text{ArH}$), 7.41 (dd, $J = 6.0, 2.8$ Hz, 1H, ArH), 7.71 (s, 1H, ArH), 8.10 – 8.03 (m, 1H, ArH), 8.37 (d, $J = 5.3$ Hz, 1H, ArH), 9.09 (s, 1H, ArH), 9.60 (s, 1H, ArNH), 9.70 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 1.1, 24.9, 31.9, 33.4, 44.0, 45.6, 56.0, 56.3, 57.5, 83.3, 104.6, 107.9, 109.7, 110.2, 113.6, 120.3, 120.9, 121.8, 126.1, 127.7, 129.8, 134.1, 135.4, 138.4, 143.9, 157.9, 159.7, 162.2, 171.5 ppm; ^{11}B NMR (96 MHz, CDCl_3) δ 33.35 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{34}\text{H}_{37}\text{N}_7\text{O}_4\text{B}$: 628.3783; found: 628.3775.

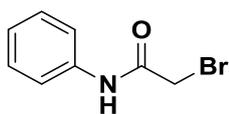
***N*-[2-([2-(Dimethylamino)ethyl](methylamino)-4-methoxy-5-{[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]amino}phenyl)-2,2,2-trifluoroacetamide (57)⁸⁰**



A flame-dried, 2-neck round-bottom flask was charged with *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (75.0 mg, 0.168 mmol, 1 equiv.), purged with Ar and dissolved in THF (4 mL). DIPEA (35.0 μL , 0.202 mmol, 1.2 equiv.) was added and the solution was cooled to -10 $^\circ\text{C}$ (acetone/ ice). A solution of trifluoroacetic anhydride (39.0 mg, 0.185 mmol, 1.1 equiv.) in THF (1 mL) was then added dropwise to the solution over 2 min. The reaction mixture was then allowed to warm to rt and stirred for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (20 mL), a saturated solution of NaHCO_3 (20 mL) and DCM (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **57** as a white foam (66 mg, 0.121 mmol, 74%).

IR (ATR) cm^{-1} : 3423 (m, N-H stretch), 3133 (m, N-H stretch), 2924 (m, C-H stretch), 2846 (m, C-H stretch), 1715 (s, C=O stretch), 1588 (m, N-H bend), 1445 (m, C-H bend), 1402 (m, C-H bend), 1340 (s, C-N stretch), 1213 (s, CF_3 stretch), 1099 (s, C-O stretch), 799 (s, C-H bend), 739 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 2.18 (t, $J = 5.9$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.22 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.72 (s, 3H, ArNCH_3), 2.98 (t, $J = 5.9$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.93 (s, 3H, ArNCH_3), 3.99 (s, 3H, ArOCH_3), 6.84 (s, 1H, ArH), 6.84 (s, 1H, ArH), 7.25 (d, $J = 5.3$ Hz, 1H, ArH), 7.32 – 7.28 (m, 2H, $2 \times \text{ArH}$), 7.46 – 7.40 (m, 1H, ArH), 7.81 (s, 1H, ArH), 8.11 – 8.02 (m, 1H, ArH), 8.41 (d, $J = 5.3$ Hz, 1H, ArH), 8.89 (s, 1H, ArH), 9.59 (s, 1H, ArNH), 11.84 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 32.9, 44.4, 45.5, 55.7, 56.1, 56.9, 105.1, 108.3, 110.3, 111.2, 113.6, 116.7 (d, $J = 288.2$ Hz), 120.2, 121.1, 122.0, 126.0, 128.0, 128.4, 135.0, 135.4, 138.4, 146.0, 154.8 (d, $J = 36.7$ Hz), 158.0, 159.6, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{27}\text{H}_{31}\text{N}_7\text{O}_2\text{F}_3$: 542.2491; found: 542.2493.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

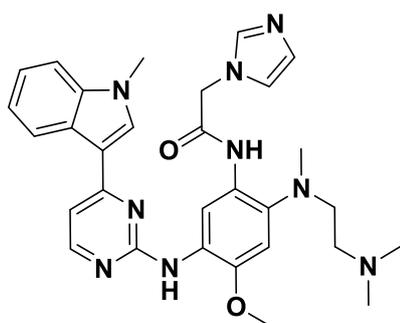
2-Bromo-*N*-phenylacetamide (61)

A flame-dried, 2-neck round-bottom flask was charged with aniline (500 mg, 5.37 mmol, 1 equiv.), purged with Ar and dissolved in DCM (15 mL). To this was added DIPEA (1.12 mL, 6.44 mmol, 1.2 equiv.) and the solution was cooled to $-40\text{ }^{\circ}\text{C}$ (acetonitrile/ CO_2).

A solution of 2-bromoacetyl bromide (1.19 g, 5.91 mmol, 1.1 equiv.) in DCM (10 mL) was then added dropwise to the solution over 5 min. The reaction mixture was then allowed to warm to $-20\text{ }^{\circ}\text{C}$ and stirred at this temperature for 2 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (40 mL), a saturated solution of NaHCO_3 (40 mL) and DCM (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM ($3 \times 40\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine (50 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 10 – 25% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **61** as a white solid (1.03 g, 4.82 mmol, 90%).

IR, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.⁸¹

Rf: 0.18 (20% EtOAc in PE); ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 4.04 (s, 2H, COCH_2Br), 7.11 – 7.06 (m, 1H, ArH), 7.36 – 7.29 (m, 2H, $2 \times$ ArH), 7.62 – 7.56 (m, 2H, $2 \times$ ArH), 10.37 (s, 1H, ArNH) ppm; ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 30.4, 119.2, 123.8, 128.8, 138.6, 164.8 ppm; HRMS-TOF MS ESI–: m/z $[\text{M}-\text{H}]^+$ calculated for $\text{C}_8\text{H}_7\text{NOBr}$: 211.9711; found: 211.9706.

***N*-[2-([2-(dimethylamino)ethyl](methyl)amino)-4-methoxy-5-{[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]amino}phenyl)-2-(1*H*-imidazol-1-yl)acetamide (67)**

A flame dried 2-neck round-bottom flask was charged with 2-(1*H*-imidazol-1-yl)acetic acid (37.0 mg, 0.292 mmol, 2 equiv.), purged with Ar and dissolved in DCM (6 mL). The resulting solution was cooled to $0\text{ }^{\circ}\text{C}$ and this was followed by the portion wise addition of 1-hydroxybenzotriazole monohydrate (45.0 mg, 0.292 mmol, 2 equiv.) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (56.0 mg, 0.292 mmol, 2 equiv.) successively over 5 min. After allowing to stir for 30 min at $0\text{ }^{\circ}\text{C}$,

a solution of *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (65.0 mg, 0.146 mmol, 1 equiv.) in DCM (4 mL) was added dropwise. The reaction mixture was stirred at $0\text{ }^{\circ}\text{C}$ for 1 h and then allowed to warm to rt and stirred for a further 12 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (30 mL), a saturated solution of NaHCO_3 (30 mL) and DCM (30 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM ($3 \times 30\text{ mL}$) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO_4 and filtered.

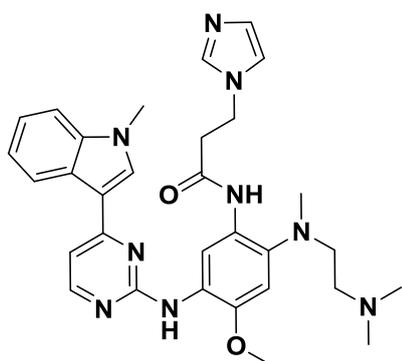
Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 10% – 25% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **67** as a light-brown solid (61.0 mg, 0.110 mmol, 75%).

Rf: 0.21 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3425 (m, N-H stretch), 3120 (m, N-H stretch), 2924 (m, C-H stretch), 2848 (m, C-H stretch), 1697 (s, C=O stretch), 1577 (m, N-H bend), 1452 (m, C-H bend), 1410 (m, C-H bend), 1342 (s, C-N stretch), 1099 (s, C-O stretch), 805 (s, C-H bend), 751 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 2.30 (t, $J = 5.7$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.34 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.65 (s, 3H, ArNCH_3), 2.92 (t, $J = 5.7$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.61 (s, 3H, ArNCH_3), 3.89 (s, 3H, ArOCH_3), 4.87 (s, 2H, COCH_2NCH), 6.77 (s, 1H, ArH), 7.07 (s, 1H, NCHCHN), 7.14 (s, 1H, NCHCHN), 7.20 (d, $J = 5.3$ Hz, 1H, ArH), 7.25 (dd, $J = 6.2, 2.7$ Hz, 1H, ArH), 7.28 (s, 1H, CH_2NCHN), 7.36 (dd, $J = 6.2, 2.7$ Hz, 1H, ArH), 7.63 (s, 1H, ArH), 7.75 (s, 1H, ArH), 8.04 (dd, $J = 6.2, 2.7$ Hz, 1H, ArH), 8.36 (d, $J = 5.3$ Hz, 1H, ArH), 8.88 (s, 1H, ArNH), 9.67 (s, 1H, ArNH) ppm;

^{13}C NMR (75 MHz, CDCl_3) δ 32.9, 44.3, 45.4, 50.4, 55.5, 56.2, 57.2, 104.5, 108.1, 109.4, 110.3, 113.4, 120.2, 120.3, 121.0, 121.9, 125.9, 128.0, 128.6, 129.7, 134.2, 135.0, 138.3, 144.6, 158.0, 159.5, 162.2, 163.7 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{30}\text{H}_{36}\text{N}_9\text{O}_2$: 554.2992; found: 554.3007.

N-(2-([2-(Dimethylamino)ethyl](methyl)amino)-4-methoxy-5-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)-3-(1H-imidazol-1-yl)propanamide (69)



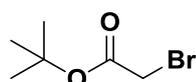
A 2-neck round-bottomed flask was charged with 1,1'-carbonyldiimidazole (201 mg, 1.25 mmol, 1.85 equiv.), purged with Ar and dissolved in DCM (10 mL). The resulting solution was then cooled to 0 °C, followed by the dropwise addition of a solution of 3-bromopropanoic acid (191 mg, 1.25 mmol, 1.85 equiv.) in DCM (2 mL). The reaction mixture was then allowed to warm to rt and stirred for 1 h before being recooled to 0 °C. A solution of N1-[2-(dimethylamino)ethyl]-5-methoxy-N1-methyl-N4-[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (300 mg, 0.673 mmol, 1 equiv.) in

DCM (6 mL) was then added dropwise to the solution and the reaction mixture was then allowed to stir for 12 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (30 mL), a saturated solution of NaHCO_3 (20 mL) and DCM (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 25% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **69** as a tan solid (241 mg, 0.424 mmol, 63%).

Rf: 0.24 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3412 (m, N-H stretch), 3114 (m, N-H stretch), 2937 (m, C-H stretch), 2819 (m, C-H stretch), 1666 (s, C=O stretch), 1577 (m, N-H bend), 1447 (m, C-H bend), 1402 (m, C-H bend), 1336 (s, C-N stretch), 1099 (s, C-O stretch), 807 (s, C-H bend), 739 (s, N-H wag);

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

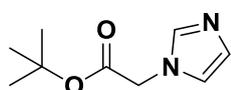
^1H NMR (300 MHz, CDCl_3) δ 2.22 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.28 (t, $J = 5.9$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.65 (s, 3H, ArNCH_3), 2.92 – 2.84 (m, 4H, $\text{NCH}_2\text{CH}_2\text{N}$ and $\text{COCH}_2\text{CH}_2\text{N}$), 3.89 (s, 3H, ArNCH_3), 3.97 (s, 3H, ArOCH_3), 4.44 (t, $J = 6.6$ Hz, 2H, $\text{COCH}_2\text{CH}_2\text{N}$), 6.78 (s, 1H, ArH), 7.03 (s, 2H, NCHCHN), 7.20 (d, $J = 5.3$ Hz, 1H, ArH), 7.32 – 7.27 (m, 2H, $2 \times \text{ArH}$), 7.41 (dd, $J = 6.2, 2.8$ Hz, 1H, ArH), 7.58 (s, 1H, CH_2NCHN), 7.73 (s, 1H, ArH), 8.16 (dd, $J = 6.2, 2.8$ Hz, 1H, ArH), 8.41 (d, $J = 5.3$ Hz, 1H, ArH), 8.81 (s, 1H, ArH), 9.61 (s, 1H, ArNH), 10.19 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 33.2, 39.0, 43.2, 44.1, 45.2, 55.9, 56.2, 57.1, 104.7, 108.1, 110.1, 110.5, 113.9, 119.2, 120.7, 121.1, 122.1, 126.0, 127.7, 129.2, 129.7, 134.2, 134.7, 137.2, 138.2, 144.7, 157.9, 159.7, 162.2, 167.1 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{31}\text{H}_{38}\text{N}_9\text{O}_2$: 568.3148; found: 568.3158.

tert-Butyl 2-bromoacetate (70)⁴⁰

A flame-dried, 2-neck round-bottom flask was charged with anhydrous *tert*-butanol (1.90 mL, 20.0 mmol, 1 equiv.), purged with Ar and dissolved in DCM (40 mL). To this was added pyridine (2.42 mL, 30.0 mmol, 1.5 equiv.) and the solution was cooled to -10 °C (acetone/ice). A solution of 2-bromoacetyl bromide (4.44 g, 22.0 mmol, 1.1 equiv.) in DCM (20 mL) was then added dropwise to the solution over 20 min. The reaction mixture was then allowed to stir at 0 °C for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then quenched and diluted with H_2O (40 mL), a saturated solution of NaHCO_3 (40 mL) and diethyl ether (50 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of diethyl ether (3×40 mL) and the combined organic layers were then washed with a solution of 1 M NaOH (30 mL), saturated solution of brine (50 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, the residue was dried under vacuum to afford compound **70** as a clear oil (2.73 g, 14.0 mmol, 70%) with no further purification required.

IR, ^1H and ^{13}C NMR data collected for this compound compared well with the reported literature values.⁴⁰

^1H NMR (300 MHz, CDCl_3) δ 1.48 (s, 9H, $(\text{CH}_3)_3\text{COC}$), 3.75 (s, 2H, COCH_2Br) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 27.8, 28.0, 83.0, 166.4 ppm; no ionization was found in MS.

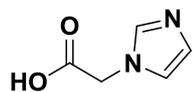
tert-Butyl 2-(1H-imidazol-1-yl)acetate (71)⁴¹

A 2-neck round-bottom flask was charged with imidazole (68.0 mg, 1.00 mmol, 1 equiv.), Cs_2CO_3 (391 mg, 1.20 mmol, 1.2 equiv.), purged with Ar and dissolved in DMF (4 mL). The resulting solution was cooled to 0 °C and to this was added a solution of *tert*-butyl 2-bromoacetate (215 mg, 1.10 mmol, 1.1 equiv.) in DMF (1 mL) dropwise over 2 min. The reaction mixture was then allowed to warm to rt and stirred for 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (30 mL) and EtOAc (20 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of EtOAc (3×20 mL) and the combined organic layers were then washed with a saturated solution of brine (3×30 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 50% – 90% EtOAc in PE. Pure fractions were evaporated to dryness to afford compound **71** as a tan solid (153 mg, 0.840 mmol, 84%).

IR, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{82, 83}

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

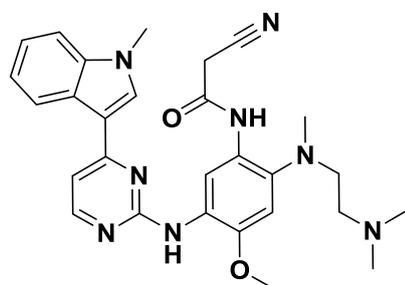
Rf: 0.17 (60% EtOAc in PE); ^1H NMR (300 MHz, CDCl_3) δ 1.45 (s, 9H, $(\text{CH}_3)_3\text{COC}$), 4.56 (s, 2H, COCH_2NCH), 6.93 (s, 1H, NCHCHN), 7.07 (s, 1H, NCHCHN), 7.47 (s, 1H, CH_2NCHN) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 28.0, 48.9, 83.3, 120.1, 129.7, 166.5 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_9\text{H}_{15}\text{N}_2\text{O}_2$: 183.1134; found: 183.1128.

2-(1H-Imidazol-1-yl)acetic acid (72)³⁶

A 2-neck round-bottomed flask was charged with *tert*-butyl 2-(1*H*-imidazol-1-yl)acetate (80.0 mg, 0.439 mmol, 1 equiv.), purged with Ar and dissolved in DCM (6 mL). This was followed by the dropwise addition of trifluoroacetic acid (2 mL) and the reaction mixture was then allowed to stir for 6 h or until complete consumption of the starting material as indicated by TLC. After removal of the solvent *in vacuo*, the residue was dried under vacuum to afford compound (HL1184 **72**) as a beige solid (49.0 mg, 0.386 mmol, 88%) with no further purification required.

IR, ^1H and ^{13}C NMR and MS data collected for this compound compared well with the reported literature values.^{84, 85, 86}

Rf: 0.08 (EtOAc); ^1H NMR (300 MHz, CD_3OD) δ 4.46 (s, 2H, COCH_2NCH), 6.86 (s, 1H, NCHCHN), 6.92 (s, 1H, NCHCHN), 8.26 (s, 1H, CH_2NCHN) ppm; ^{13}C NMR (75 MHz, CD_3OD) δ 50.6, 120.6, 124.6, 137.7, 169.4 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{12}\text{H}_9\text{N}_3\text{Cl}$: 230.0485; found: 230.0480.

2-Cyano-N-(2-{2-(dimethylamino)ethyl}(methyl)amino}-4-methoxy-5-{[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino}phenyl)acetamide (74)

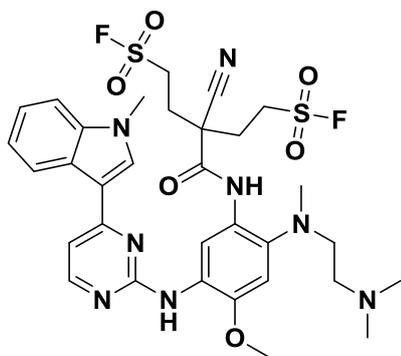
A flame dried 2-neck round-bottom flask was charged with 2-cyanoacetic acid (85.0 mg, 2.00 mmol, 2 equiv.), purged with Ar and dissolved in DCM (12 mL). The resulting solution was cooled to 0 °C and this was followed by the portion wise addition of 1-hydroxybenzotriazole monohydrate (153 mg, 2.00 mmol, 2 equiv.) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (192 mg, 2.00 mmol, 2 equiv.) successively over 5 min. After

allowing to stir for 30 min at 0 °C, a solution of *N*1-[2-(dimethylamino)ethyl]-5-methoxy-*N*1-methyl-*N*4-[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]benzene-1,2,4-triamine (445 mg, 1.00 mmol, 1 equiv.) in DCM (8 mL) was added dropwise. The reaction mixture was stirred at 0 °C for 1 h and then allowed to warm to rt and stirred for a further 4 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then diluted with H_2O (30 mL), a saturated solution of NaHCO_3 (30 mL) and DCM (40 mL) and the organic layer was separated. The aqueous layer was extracted with aliquots of DCM (3×30 mL) and the combined organic layers were then washed with a saturated solution of brine (40 mL), dried over MgSO_4 and filtered. After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5% – 10% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **74** as a light-brown solid (466 mg, 0.909 mmol, 91%).

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

Rf: 0.48 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3419 (m, N-H stretch), 3120 (m, N-H stretch), 2939 (m, C-H stretch), 2819 (m, C-H stretch), 2253 (w, $\text{C}\equiv\text{N}$ stretch), 1687 (s, $\text{C}=\text{O}$ stretch), 1577 (m, N-H bend), 1450 (m, C-H bend), 1402 (m, C-H bend), 1345 (s, C-N stretch), 1099 (s, C-O stretch), 807 (s, C-H bend), 739 (s, N-H wag); ^1H NMR (600 MHz, CDCl_3) δ 2.26 (t, $J = 5.4$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.30 (s, 6H, $\text{CH}_2\text{N}(\text{CH}_3)_2$), 2.70 (s, 3H, ArNCH_3), 2.88 (t, $J = 5.4$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.58 (s, 2H, COCH_2CN), 3.89 (s, 3H, ArNCH_3), 4.06 (s, 3H, ArOCH_3), 6.79 (s, 1H, ArH), 7.23 (d, $J = 5.3$ Hz, 1H, ArH), 7.27 (d, $J = 4.0$ Hz, 1H, ArH), 7.33 – 7.29 (m, 1H, ArH), 7.42 (dd, $J = 6.2, 2.8$ Hz, 1H, ArH), 7.76 (s, 1H, ArH), 8.05 (dd, $J = 6.2, 2.8$ Hz, 1H, ArH), 8.38 (d, $J = 5.3$ Hz, 1H, ArH), 9.02 (s, 1H, ArH), 9.66 (s, 1H, ArNH), 10.57 (s, 1H, ArNH) ppm; ^{13}C NMR (75 MHz, CDCl_3) δ 27.2, 33.43, 44.1, 45.5, 56.2, 56.4, 57.4, 104.6, 108.2, 109.6, 110.4, 113.5, 115.0, 120.2, 121.0, 121.9, 126.0, 128.1, 128.8, 134.5, 135.1, 138.4, 144.8, 158.0, 158.5, 159.6, 162.2 ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{28}\text{H}_{33}\text{N}_8\text{O}_2$: 513.2726; found: 513.2726.

3-Cyano-3-[(2-[[2-(dimethylamino)ethyl](methyl)amino]-4-methoxy-5-[[4-(1-methyl-1H-indol-3-yl)pyrimidin-2-yl]amino]phenyl]carbamoyl]pentane-1,5-disulfonyl difluoride (76)⁴⁸



A 2-neck round-bottom flask was charged with 2-cyano-*N*-(2-[[2-(dimethylamino)ethyl](methyl)amino]-4-methoxy-5-[[4-(1-methyl-1*H*-indol-3-yl)pyrimidin-2-yl]amino]phenyl)acetamide (50.0 mg, 97.5 μmol , 1 equiv.) and potassium carbonate (27.0 mg, 0.195 mmol, 2 equiv.) and purged with Ar. To this was added acetonitrile (8 mL) and the suspension was cooled to -10 $^{\circ}\text{C}$ (acetone/ice). A solution of ethenesulfonyl fluoride (20.0 mg, 0.180 mmol, 1.85 equiv.) in acetonitrile (2 mL) was then added dropwise to the solution over 2 min. The reaction mixture was then allowed to warm to rt and

stirred for 8 h or until complete consumption of the starting material as indicated by TLC. The reaction mixture was then filtered through a plug of celite and washed with EtOAc (30 mL). After removal of the solvent *in vacuo*, purification of the crude product was achieved using flash column chromatography with elution gradient of 5 – 12.5% MeOH in DCM. Pure fractions were evaporated to dryness to afford compound **76** as an off-white foam (51.0 mg, 69.2 μmol , 71%).

Rf: 0.61 (10% MeOH in DCM); IR (ATR) cm^{-1} : 3398 (m, N-H stretch), 3120 (m, N-H stretch), 2935 (m, C-H stretch), 2236 (w, $\text{C}\equiv\text{N}$ stretch), 1682 (s, $\text{C}=\text{O}$ stretch), 1579 (m, N-H bend), 1450 (m, C-H bend), 1404 (m, C-H bend), 1335 (s, C-N stretch), 1198 (s, $\text{S}=\text{O}$ stretch), 1099 (s, C-O stretch), 1033 (s, $\text{S}=\text{O}$ stretch), 809 (s, C-H bend), 743 (s, N-H wag); ^1H NMR (300 MHz, CDCl_3) δ 2.20 (t, $J = 5.4$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 2.41 – 2.28 (m, 8H, $\text{CH}_2\text{N}(\text{CH}_3)_2$ and $\text{CCH}_2\text{CH}_2\text{S}$), 2.59 (m [app. td], $J = 13.1, 4.5$ Hz, 2H, $\text{CCH}_2\text{CH}_2\text{S}$), 2.70 (s, 3H, ArNCH_3), 2.99 (t, $J = 5.4$ Hz, 2H, $\text{NCH}_2\text{CH}_2\text{N}$), 3.41 – 3.27 (m, 2H, $\text{CCH}_2\text{CH}_2\text{S}$), 3.60 – 3.47 (m, 2H, $\text{CCH}_2\text{CH}_2\text{S}$), 3.88 (s, 3H, ArNCH_3), 3.92 (s, 3H, ArOCH_3), 7.06 (d, $J = 5.3$ Hz, 1H, ArH), 7.37 – 7.27 (m, 2H, $2 \times \text{ArH}$), 7.44 – 7.39 (m, 1H, ArH), 7.77 (s, 1H, ArH), 8.01 (s, 1H, ArH), 8.34 (d, $J = 5.3$ Hz, 1H, ArH), 8.43 (dd, $J = 6.4, 2.4$ Hz, 1H, ArH), 9.23 (s, 1H, ArNH), 10.91 (s, 1H, ArNH) ppm;

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

^{13}C NMR (75 MHz, CDCl_3) δ 31.3, 33.4, 45.0, 45.4, 46.2, 46.4, 46.7, 55.7, 56.1, 58.0, 104.4, 108.3, 110.6, 113.9, 114.0, 117.8, 121.4, 121.8, 122.5, 126.0, 127.5, 128.5, 132.2, 136.3, 138.3, 146.9, 157.5, 159.6, 161.6, 162.4 ppm; ^{19}F NMR (282 MHz, CDCl_3) δ 53.35 (t, $J = 4.8$ Hz) ppm; HRMS-TOF MS ESI+: m/z $[\text{M}+\text{H}]^+$ calculated for $\text{C}_{32}\text{H}_{39}\text{N}_8\text{O}_6\text{S}_2\text{F}_2$: 733.2402; found: 733.2407.

7.9 References

1. K. S. Thress, C. P. Paweletz, E. Felip, B. C. Cho, D. Stetson, B. Dougherty, Z. Lai, A. Markovets, A. Vivancos, Y. Kuang, D. Ercan, S. E. Matthews, M. Cantarini, J. C. Barrett, P. A. Jänne and G. R. Oxnard, *Nature Medicine*, 2015, **21**, 1-5.
2. P. A. Jackson, J. C. Widen, D. A. Harki and K. M. Brummond, *Journal of Medicinal Chemistry*, 2017, **60**, 839-885.
3. G. Dodson and A. Wlodawer, *Trends in Biochemical Sciences*, 1998, **23**, 347-352.
4. J. Phan, A. Zdanov, A. G. Evdokimov, J. E. Tropea, H. K. Peters, R. B. Kapust, M. Li, A. Wlodawer and D. S. Waugh, *Journal of Biological Chemistry*, 2002, **277**, 50564-50572.
5. A. R. Buller and C. A. Townsend, *Proceedings of the National Academy of Sciences of the United States of America*, 2013, **110**, E653-E661.
6. L. Simón and J. M. Goodman, *Journal of Organic Chemistry*, 2010, **75**, 1831-1840.
7. J. C. Powers, J. L. Asgian, Ö. D. Ekici and K. E. James, *Chemical Reviews*, 2002, **102**, 4639-4750.
8. A. Narayanan and L. H. Jones, *Chemical Science*, 2015, **6**, 2650-2659.
9. O. O. Fadeyi, L. R. Hoth, C. Choi, X. Feng, A. Gopalsamy, E. C. Hett, R. E. Kyne, R. P. Robinson and L. H. Jones, *ACS Chemical Biology*, 2017, **12**, 2015-2020.
10. F. F. Fleming, L. Yao, P. C. Ravikumar, L. Funk and B. C. Shook, *Journal of Medicinal Chemistry*, 2010, **53**, 7902-7917.
11. A. Berteotti, F. Vacondio, A. Lodola, M. Bassi, C. Silva, M. Mor and A. Cavalli, *ACS Medicinal Chemistry Letters*, 2014, **5**, 501-505.
12. L.-M. Ni and J. C. Powers, *Bioorganic & Medicinal Chemistry*, 1998, **6**, 1767-1773.
13. J. Oleksyszyn and J. C. Powers, *Biochemical and Biophysical Research Communications*, 1989, **161**, 143-149.
14. S. J. Baker, C. Z. Ding, T. Akama, Y.-K. Zhang, V. Hernandez and Y. Xia, *Future Medicinal Chemistry*, 2009, **1**, 1275-1288.
15. H. DeFrancesco, J. Dudley and A. Coca, in *Boron Reagents in Synthesis*, American Chemical Society, 2016, vol. 1236, ch. 1, pp. 1-25.
16. B. C. Das, P. Thapa, R. Karki, C. Schinke, S. Das, S. Kambhampati, S. K. Banerjee, P. V. Veldhuizen, A. Verma, L. M. Weiss and T. Evans, *Future Medicinal Chemistry*, 2013, **5**, 653-676.
17. D. A. Shannon and E. Weerapana, *Current Opinion in Chemical Biology*, 2015, **24**, 18-26.
18. A. Paramore and S. Frantz, *Nature Reviews Drug Discovery*, 2003, **2**, 611-612.
19. H. S. Ban, T. Usui, W. Nabeyama, H. Morita, K. Fukuzawa and H. Nakamura, *Organic & Biomolecular Chemistry*, 2009, **7**, 4415-4427.
20. H. Nakamura, R. Horikoshi, T. Usui and H. S. Ban, *MedChemComm*, 2010, **1**, 282-286.
21. C. B. Kelly, M. A. Mercadante and N. E. Leadbeater, *Chemical Communications*, 2013, **49**, 11133-11148.
22. R. A. Ward, M. J. Anderton, S. Ashton, P. A. Bethel, M. Box, S. Butterworth, N. Colclough, C. G. Chorley, C. Chuaqui, D. A. E. Cross, L. A. Dakin, J. É. Debreczeni, C. Eberlein, M. R. V. Finlay, G. B. Hill, M. Grist, T. C. M. Klinowska, C. Lane, S. Martin, J. P. Orme, P. Smith, F. Wang and M. J. Waring, *Journal of Medicinal Chemistry*, 2013, **56**, 7025-7048.
23. M. R. V. Finlay, M. Anderton, S. Ashton, P. Ballard, P. A. Bethel, M. R. Box, R. H. Bradbury, S. J. Brown, S. Butterworth, A. Campbell, C. Chorley, N. Colclough, D. A. E. Cross, G. S. Currie, M. Grist, L. Hassall, G. B. Hill, D. James, M. James, P. Kemmitt, T. Klinowska, G. Lamont, S. G. Lamont, N. Martin, H. L. McFarland, M. J. Mellor, J. P. Orme, D. Perkins, P. Perkins, G. Richmond, P. Smith, R. A. Ward, M. J. Waring, D. Whittaker, S. Wells and G. L. Wrigley, *Journal of Medicinal Chemistry*, 2014, **57**, 8249-8267.
24. D. Saha, R. Ghosh and A. Sarkar, *Tetrahedron*, 2013, **69**, 3951-3960.
25. J. A. Hyatt and J. J. Krutak, *Journal of Organic Chemistry*, 1977, **42**, 169-170.
26. A. R. Moosavi-Zare, M. A. Zolfigol, E. Noroozizadeh, M. Zarei, R. Karamian and M. Asadbegy, *Journal of Molecular Catalysis A: Chemical*, 2016, **425**, 217-228.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

27. C. Grison, P. Coutrot, C. Comoy, L. Balas, S. Joliez, G. Lavecchia, P. Oligier, B. Penverne, V. Serre and G. Hervé, *European Journal of Medicinal Chemistry*, 2004, **39**, 333-344.
28. H. Krawczyk, *Synthetic Communications*, 1997, **27**, 3151-3161.
29. S. R. Houghton, J. Melton, J. Fortunak, D. H. Brown Ripin and C. N. Boddy, *Tetrahedron*, 2010, **66**, 8137-8144.
30. M. D. M. Gray and D. J. H. Smith, *Tetrahedron Letters*, 1980, **21**, 859-860.
31. G. A. Molander and S. A. McKee, *Organic Letters*, 2011, **13**, 4684-4687.
32. J. Sun, M. T. Perfetti and W. L. Santos, *Journal of Organic Chemistry*, 2011, **76**, 3571-3575.
33. S. Hermanek, *Chemical Reviews*, 1992, **92**, 325-362.
34. B. W. Glasspoole, K. Ghazati, J. W. Moir and C. M. Crudden, *Chemical Communications*, 2012, **48**, 1230-1232.
35. R. K. Haynes, S. C. Vonwiller and M. R. Luderer, in *Encyclopedia of Reagents for Organic Synthesis*, 2006, DOI: doi:10.1002/047084289X.rt064.pub2.
36. B. F. Lundt, N. L. Johansen, A. Vølund and J. Markussen, *International Journal of Peptide and Protein Research*, 1978, **12**, 258-268.
37. T. G. Elford, S. Nave, R. P. Sonawane and V. K. Aggarwal, *Journal of the American Chemical Society*, 2011, **133**, 16798-16801.
38. A. C. Flick, H. X. Ding, C. A. Leverett, R. E. Kyne, K. K. C. Liu, S. J. Fink and C. J. O'Donnell, *Journal of Medicinal Chemistry*, 2017, **60**, 6480-6515.
39. H. Walba and R. W. Isensee, *Journal of Organic Chemistry*, 1961, **26**, 2789-2791.
40. J. Pospíšil and M. Potáček, *Tetrahedron*, 2007, **63**, 337-346.
41. M. Milen, A. Grün, E. Bálint, A. Dancsó and G. Keglevich, *Synthetic Communications*, 2010, **40**, 2291-2301.
42. R. Mondal and A. K. Mallik, *Organic Preparations and Procedures International*, 2014, **46**, 391-434.
43. J. Dong, L. Krasnova, M. G. Finn and K. B. Sharpless, *Angewandte Chemie International Edition*, 2014, **53**, 9430-9448.
44. M. J. Minch, *Concepts in Magnetic Resonance*, 1994, **6**, 41-56.
45. H. K. Hall, *Journal of the American Chemical Society*, 1957, **79**, 5441-5444.
46. F. S. Prout, A. A. Abdel-Latif and M. R. Kamal, *Journal of Chemical & Engineering Data*, 1963, **8**, 597-599.
47. D. E. Pearson and C. A. Buehler, *Chemical Reviews*, 1974, **74**, 45-86.
48. J. J. Krutak, R. D. Burpitt, W. H. Moore and J. A. Hyatt, *Journal of Organic Chemistry*, 1979, **44**, 3847-3858.
49. I. Fleming, *Molecular Orbitals and Organic Chemical Reactions, Reference Edition*, 2010.
50. H. Mukherjee, J. Debreczeni, J. Breed, S. Tentarelli, B. Aquila, J. E. Dowling, A. Whitty and N. P. Grimster, *Organic & Biomolecular Chemistry*, 2017, **15**, 9685-9695.
51. D. H. Lee and A. L. Goldberg, *Trends in Cell Biology*, 1998, **8**, 397-403.
52. J. F. King and T. Durst, *Journal of the American Chemical Society*, 1964, **86**, 287-288.
53. A. M. Gold, in *Methods in Enzymology*, Academic Press, 1967, vol. 11, pp. 706-711.
54. J. Graton, J.-Y. Le Questel, P. Maxwell and P. Popelier, *Journal of Chemical Information and Modeling*, 2016, **56**, 322-334.
55. R. N. V. Krishna Deepak and R. Sankaramakrishnan, *Biochemistry*, 2016, **55**, 3774-3783.
56. A. L. Ringer, A. Senenko and C. D. Sherrill, *Protein Science*, 2007, **16**, 2216-2223.
57. F. Zhou, R. Liu, P. Li and H. Zhang, *New Journal of Chemistry*, 2015, **39**, 1611-1618.
58. S. K. Singh and A. Das, *Physical Chemistry Chemical Physics*, 2015, **17**, 9596-9612.
59. H. Park, H.-Y. Jung, S. Mah and S. Hong, *Angewandte Chemie International Edition*, 2017, **56**, 7634-7638.
60. T. Grabe, J. Lategahn and D. Rauh, *ACS Medicinal Chemistry Letters*, 2018, **9**, 779-782.
61. M. Toure, O. Chuzel and J.-L. Parrain, *Dalton Transactions*, 2015, **44**, 7139-7143.
62. L. Liu, K. Sun, L. Su, J. Dong, L. Cheng, X. Zhu, C.-T. Au, Y. Zhou and S.-F. Yin, *Organic Letters*, 2018, **20**, 4023-4027.
63. S.-H. Chung, T.-J. Lin, Q.-Y. Hu, C.-H. Tsai and P.-S. Pan, *Molecules*, 2013, **18**, 12346.
64. P. K. Chinthakindi, K. B. Govender, A. S. Kumar, H. G. Kruger, T. Govender, T. Naicker and P. I. Arvidsson, *Organic Letters*, 2017, **19**, 480-483.
65. A. T. Davies, J. M. Curto, S. W. Bagley and M. C. Willis, *Chemical Science*, 2017, **8**, 1233-1237.
66. A. L. Tribby, I. Rodríguez, S. Shariffudin and N. D. Ball, *Journal of Organic Chemistry*, 2017, **82**, 2294-2299.
67. T. Guo, G. Meng, X. Zhan, Q. Yang, T. Ma, L. Xu, K. B. Sharpless and J. Dong, *Angewandte Chemie International Edition*, 2018, **57**, 2605-2610.
68. H. Hemit, J. Yong, L. Nie and H. A. Aisa, *Synthetic Communications*, 2010, **40**, 973-979.
69. J. Jaśkowska and P. Kowalski, *Journal of Heterocyclic Chemistry*, 2008, **45**, 1371-1375.

Chapter 7 – Osimertinib-derived Inhibitors Targeting the Mutated Serine Residue Ser797

70. C. P. Bergstrom, C. P. Sloan, W.-Y. Lau, D. W. Smith, M. Zheng, S. B. Hansel, C. T. Polson, J. A. Corsa, D. M. Barten, K. M. Felsenstein and S. B. Roberts, *Bioorganic & Medicinal Chemistry Letters*, 2008, **18**, 464-468.
71. S. Arjan H. G., W. Wei Sein and L. Johan, *European Journal of Organic Chemistry*, 2003, **2003**, 4664-4678.
72. S. Sano, E. Kujime, Y. Takemoto, M. Shiro and Y. Nagao, *Chemical and Pharmaceutical Bulletin*, 2005, **53**, 131-134.
73. G. G. Rajeshwaran, M. Nandakumar, R. Sureshbabu and A. K. Mohanakrishnan, *Organic Letters*, 2011, **13**, 1270-1273.
74. H. Qi, X. Li and J. Xu, *Organic & Biomolecular Chemistry*, 2011, **9**, 2702-2714.
75. A. Perez-Anes, G. Spataro, Y. Coppel, C. Moog, M. Blanzat, C.-O. Turrin, A.-M. Caminade, I. Rico-Lattes and J.-P. Majoral, *Organic & Biomolecular Chemistry*, 2009, **7**, 3491-3498.
76. D. Brandt, A. Dittoo, V. Bellosta and J. Cossy, *Organic Letters*, 2015, **17**, 816-819.
77. F. Scaravelli, S. Bacchi, L. Massari, O. Curcuruto, P. Westerduin and W. Maton, *Tetrahedron Letters*, 2010, **51**, 5154-5156.
78. L. A. Bateman, T. B. Nguyen, A. M. Roberts, D. K. Miyamoto, W. M. Ku, T. R. Huffman, Y. Petri, M. J. Heslin, C. M. Contreras, C. F. Skibola, J. A. Olzmann and D. K. Nomura, *Chemical Communications*, 2017, **53**, 7234-7237.
79. M. Gao, S. B. Thorpe and W. L. Santos, *Organic Letters*, 2009, **11**, 3478-3481.
80. B. Kim, A. J. Chinn, D. R. Fandrick, C. H. Senanayake, R. A. Singer and S. J. Miller, *Journal of the American Chemical Society*, 2016, **138**, 7939-7945.
81. W. A. Loughlin, I. D. Jenkins, N. D. Karis, S. S. Schweiker and P. C. Healy, *European Journal of Medicinal Chemistry*, 2016, **111**, 1-14.
82. P. Zaderenko, M. S. Gil, P. Ballesteros and S. Cerdan, *Journal of Organic Chemistry*, 1994, **59**, 6268-6273.
83. J. F. Jensen, W.-L. Kasper and M. Morten, *European Journal of Organic Chemistry*, 2008, **2008**, 3785-3797.
84. P. López, P. Zaderenko, J. L. Balcazar, I. Fonseca, F. H. Cano and P. Ballesteros, *Journal of Molecular Structure: THEOCHEM*, 1996, **377**, 105-112.
85. G. Yang, H. Du, J. Liu, Z. Zhou, X. Hu and Z. Zhang, *Green Chemistry*, 2017, **19**, 675-681.
86. L. Men, Z. Pi, Y. Zhou, Y. Liu, M. Wei, F. Song and Z. Liu, *RSC Advances*, 2017, **7**, 16494-16504.

Chapter 8

Project Summary and Concluding Remarks

8.1 The Future of EGFR as Therapeutic Target

From the revelation of its existence and mechanistic functions in the 1950s, to the awarding of the 1986 Nobel Prize in Physiology and Medicine to Stanley Cohen for its discovery, EGFR has remained a relevant target in cancer therapy for 36 years.¹ The FDA approval of gefitinib as the first TKI for treatment of NSCLC in 2003 marked a major triumph in the development of targeted therapeutics against cancer.² However, this victory was short-lived. The past 16 years has seen the development of four generations of EGFR inhibitors attempting to overcome the emergence of multiple, major resistance mechanisms which are brought about by point mutations within the enzyme.³ During this time, researchers have been forced to re-evaluate and implement numerous innovative strategies, both in the identification and targeting of EGFR mutant variants. This includes the discovery of novel chemical structures, binding modes and means of inhibition. This is most celebrated in the success story that is the development of osimertinib, the first inhibitor to effectively overcome resistance mediated through the most prevalent mechanism - the EGFR-L858R/T790M double mutation.⁴

Yet, four years since its emergence in 2015, resistance to 3rd generation inhibitors conferred by the EGFR-C797S mutation remains unanswered.⁵ To date, no major breakthroughs have been achieved in subduing the clinically relevant mutant variant. Furthermore, the EGFR-G724S mutation was recently identified for the first time and shown to limit the activity of 3rd generation inhibitors both in vitro and in vivo.⁶ The drought in successful strategies targeting 3rd generation inhibitor resistance, as well as the emergence of novel contributing mutations, highlights our lack of knowledge and understanding in this current wave of resistance mechanisms. While sombre, this describes the present situation of EGFR inhibition.

Owing to the inherent relationship between receptor overexpression and cancer, EGFR will remain a therapeutic target in combating various forms of cancer for the foreseeable future. Consequently, researchers must examine the successes of the past to overcome the obstacles of the future. As was learnt with the emergence of the EGFR-T790M mutation over a decade ago, simply trying to improve on existing inhibitors does not necessarily counteract new mechanisms of acquired drug resistance. With the complexity of multimutated forms of EGFR exponentially increased, researchers find themselves in a similar but more demanding and challenging situation. Therefore, the development of novel therapeutic strategies and innovative approaches, through collaboration between academia and industry, will be required to overcome mutant forms of the enzyme which mediate resistance toward 3rd generation inhibitors. This was the intention of this research project.

The research efforts described within this thesis focussed on overcoming EGFR-C797S-mediated resistance through development of both highly potent reversible inhibitors and re-establishment of irreversible inhibition. This was undertaken by attempting covalent modification of previously untargeted amino acid residues utilising novel electrophiles. In the next section, we will examine the positive and negative outcomes of each strategy employed, which will be followed by the prospective application of these approaches in the future of EGFR inhibition.

8.2 Summary, Outcomes and Future Applications of this Project

In Chapter 4, we implemented the use of traditional reversible and acrylamide-containing irreversible inhibitors, synthesising a total of 26 final compounds. For the reversible analogues, successful biochemical inhibition of the clinically relevant EGFR-L858R/T790M/C797S mutant variant was reflected in the highly potent picomolar IC_{50} values of these compounds. Unfortunately, this efficacy was not maintained in the cellular assays, presumably due to intolerance of the glycol chain moiety within the cell. However, with careful selection of a cellularly tolerated glycol chain group, and optimisation of the relevant pharmacokinetic parameters, we believe inclusion of this functional group will be beneficial in creating favourable reversible interactions within the enzyme. Therefore, it is our belief that the design and synthesis of highly potent reversible inhibitors could be a viable strategy in overcoming resistance mediated by the EGFR-C797S mutation.

For the irreversible counterparts, similar activity within the low nanomolar to picomolar range was observed against the EGFR-L858R/T790M double mutant biochemically, while a decrease in the respective EC_{50} values of these compounds was observed in a cellular context. Irreversible inhibition was confirmed through procurement of a crystal structure of compound **77** complexed with the cSrc-T338M/S345C surrogate. Additionally, this provided valuable insight and reinforcement for our postulates regarding the optimal position and substitution pattern of the glycol chain moiety for favourable reversible interactions with the glycine rich loop. While the effectiveness of osimertinib against the EGFR-L858R/T790M double mutant has rendered its therapeutic targeting a secondary objective, the development of more potent and selective inhibitors of this aberrant enzyme remains a continual research endeavour. Thereby, the aforementioned work has led to the submitting of a manuscript for publication. Lastly, the most efficacious inhibitors synthesised in this chapter were submitted for evaluation as dual EGFR/HER2 inhibitory agents. The promising activity profiles exhibited by these compounds in cellular-based assays laid the foundation for a further research platform, which remains ongoing.

Chapter 5 witnessed a departure from the conventional methods of EGFR inhibition. Seeking to re-establish covalent inhibition in the EGFR-C797S mutant, we envisaged the use of 1,4-dicarbonyl electrophiles to potentially undergo Paal-Knorr pyrrole formation with the EGFR catalytic Lys745 residue. This included synthesis of 1,4-diketone-, 1,4-dicarboxylic acid-, 1,4-diester-, 1,4-diol- and 1,4-dialkene-containing inhibitors and a gefitinib-based heterocyclic driving group. This amounted to 8 final compounds which were submitted for biochemical and cellular evaluation. While covalent mass spectrometry revealed a reversible binding mode for these inhibitors, several interesting activity profiles warrants further investigation into these novel electrophiles, with the most notable example being the 1,4-dialkene functional group. Furthermore, we believe that the inclusion of an appropriate solubilising group will dramatically improve the poor EC_{50} values displayed by these compounds.

Building on the premise of the preceding chapter, we continued with our study of the reversible/irreversible interactive capabilities of numerous functional groups with the EGFR catalytic Lys745 residue in Chapter 6. An arduous synthetic route afforded 10 final compounds functionalised at the *N*-indole position with various protecting groups, nitriles, acrylamide and sulfonyl fluoride moieties. In addition, an optimised route towards the osimertinib-derived heterocyclic scaffold was established.

Chapter 8 – Project Summary and Concluding Remarks

Unfortunately, no irreversible inhibition of the EGFR catalytic Lys745 residue was found to take place, as demonstrated by covalent mass spectrometry experiments. While covalent labelling of this residue was achieved by Taunton and co-workers, effective irreversible inhibition of EGFR through this mechanism remains a challenging endeavour.⁷ Rationally targeting nucleophilic lysine residues in a therapeutic context is currently in its infancy, with the relatively few available reported examples a testament to this.⁸ With a greater selection as these examples increase, and improved knowledge of optimal lysine-selective electrophiles and the required reactivity thereof, lysine-targeting covalent inhibition will become more prevalent and considered earlier in the drug discovery process. Nevertheless, we still maintain that covalent bond formation with the catalytic Lys745 residue remains a potentially realizable strategy in overcoming resistance conferred by the EGFR-C797S mutation. Efforts towards this goal remains ongoing in our research group.

The final approach implemented in Chapter 7 aimed to irreversibly inhibit the mutated Ser797 residue present in the EGFR-C797S mutant enzyme. Based on the synthetic achievements of the previous chapter, we utilised an osimertinib-derived driving group derivatised with various reactive electrophiles and functional groups capable of interacting irreversibly or reversibly with the Ser797 residue respectively. Wishing to emulate the catalytic triad in serine proteases, we also envisaged the use of dual-warhead-containing compounds. This amounted to a total of 14 final compounds which included nitrogen-, boron-, phosphorous-, sulphur- and fluorine-containing functionalities. While none of these compounds were able to undergo covalent bond formation with the mutant Ser797 residue, a number of encouraging trends and results were obtained from the biochemical and cellular evaluation. This included the enhanced stability of the vinyl sulfonyl fluoride warhead over its aliphatic counterpart, rationalised to stem from the prevailing mechanism of hydrolysis. Additionally, the favourable selectivity and efficacy offered by the imidazole- and nitrile-bearing inhibitors between the double and triple mutant form of the enzyme was ascribed to the sulphur- π interaction capabilities of the heteroaromatic imidazole ring. These attractive features, coupled with the novelty of these underutilised moieties, emphasises their prospective development and use in combating the EGFR-C797S mutation. With serine residues being notoriously non-nucleophilic, we remain unsure as to whether covalent bond formation with the mutated Ser797 residue is indeed possible. However, further investigation into the envisioned dual warhead strategy, which allows for activation and simultaneous irreversible inhibition of the unreactive residue. This remains a main priority in the activities of our research group.

To conclude, an extensive amount of synthetic methodology was developed and implemented in the synthesis of the 48 final compounds presented in this thesis. Most notable is the discovery of a novel and convenient one-pot transformation of anilines to their corresponding acetamide- or propanamide-functionalised imidazoles, utilising 1,1'-carbonyldiimidazole and the respective α -, β -halo-carboxylic acid. We intend on examining the scope of this reaction, with the purpose of publishing the compiled parameters. It is our hope that this synthetic sequence, along with the synthesis of the numerous electrophilic fragments and heterocycles furnished throughout this research project, will provide a platform for future drug candidates in the successful inhibition of EGFR.

8.3 References

1. R. Roskoski, *Pharmacological Research*, 2014, **79**, 34-74.
2. R. S. Herbst, M. Fukuoka and J. Baselga, *Nature Reviews Cancer*, 2004, **4**, 979-987.
3. K. S. Bhullar, N. O. Lagarón, E. M. McGowan, I. Parmar, A. Jha, B. P. Hubbard and H. P. V. Rupasinghe, *Molecular Cancer*, 2018, **17**, 17-48.
4. M. R. V. Finlay, M. Anderton, S. Ashton, P. Ballard, P. A. Bethel, M. R. Box, R. H. Bradbury, S. J. Brown, S. Butterworth, A. Campbell, C. Chorley, N. Colclough, D. A. E. Cross, G. S. Currie, M. Grist, L. Hassall, G. B. Hill, D. James, M. James, P. Kemmitt, T. Klinowska, G. Lamont, S. G. Lamont, N. Martin, H. L. McFarland, M. J. Mellor, J. P. Orme, D. Perkins, P. Perkins, G. Richmond, P. Smith, R. A. Ward, M. J. Waring, D. Whittaker, S. Wells and G. L. Wrigley, *Journal of Medicinal Chemistry*, 2014, **57**, 8249-8267.
5. K. S. Thress, C. P. Paweletz, E. Felip, B. C. Cho, D. Stetson, B. Dougherty, Z. Lai, A. Markovets, A. Vivancos, Y. Kuang, D. Ercan, S. E. Matthews, M. Cantarini, J. C. Barrett, P. A. Jänne and G. R. Oxnard, *Nature Medicine*, 2015, **21**, 1-5.
6. J. Fassunke, F. Müller, M. Keul, S. Michels, M. A. Dammert, A. Schmitt, D. Plenker, J. Lategahn, C. Heydt, J. Brägelmann, H. L. Tumbrink, Y. Alber, S. Klein, A. Heimsoeth, I. Dahmen, R. N. Fischer, M. Scheffler, M. A. Ihle, V. Priesner, A. H. Scheel, S. Wagener, A. Kron, K. Frank, K. Garbert, T. Persigehl, M. Püsken, S. Haneder, B. Schaaf, E. Rodermann, W. Engel-Riedel, E. Felip, E. F. Smit, S. Merkelbach-Bruse, H. C. Reinhardt, S. M. Kast, J. Wolf, D. Rauh, R. Büttner and M. L. Sos, *Nature Communications*, 2018, **9**, 4655.
7. Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame and J. Taunton, *Journal of the American Chemical Society*, 2017, **139**, 680-685.
8. J. Pettinger, K. Jones and M. D. Cheeseman, *Angewandte Chemie International Edition*, 2017, **56**, 15200-15209.