

Design and synthesis of novel antimalarial agents

by
Josephus Jacobus de Jager

*Thesis presented in fulfilment of the requirements for the degree of
Master of Science in the Faculty of Science at Stellenbosch University*



Supervisor: Dr Stephen C Pelly
Co-supervisor: Prof Willem AL Otterlo

"
"
"
"
"
"

.....F gego dgt"4236

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 sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2014

Copyright © 2014 Stellenbosch University

All rights reserved

Abstract

Malaria is a pestilent disease associated with massive socioeconomic burden of sub-Saharan Africa. This disease is caused by a blood infection of the single cellular parasite of the *Plasmodium* genus. Two enzymes of this parasite have been identified to be essential to the survival of this parasite, notably Spermidine Synthase and Protein Farnesyltransferase. The goal of this dissertation was to search for and synthesise novel inhibitors of these two enzymes with a strong focus towards understanding their structure/activity relationships.

To achieve the first goal, molecular modelling was employed. An in-depth discussion is presented to describe the underlying principles relevant to this branch of computational chemistry. This ensures that the experiments using these methods are set-up correctly and results are interpreted within context. Two virtual high-throughput screens were then performed using prepared crystallographic structures of Spermidine Synthase. The first was pharmacophore based method and the second based on LibDock. The database used, containing 7.1 million compounds, was filtered using a custom developed tool prior to screening. Finally, CDOCKER was then used to investigate the activity of potential hit compounds.

Spermidine Synthase has a natural affinity for adenosine and this trait was exploited by derivatising analogues to synthesise potential inhibitors of the enzyme. This was to be achieved by the incorporation of both electrophilic and nucleophilic moieties at selected positions, including the use of a high yielding Mitsunobu reaction. A number of additional residues were then synthesised and joined to the adenosine which were proposed to increase the active site occupancy and increase affinity to the enzyme.

For the second enzyme targeted for inhibition, Protein Farnesyltransferase, indole was used as a starting scaffold to synthesise potential hits *de novo*. It was aimed to derivatise the indole at the *N* and 3' positions. The crystal structure of one of the intermediates was published. Furthermore, a synthetic sequence which culminated in a palladium catalysed Suzuki coupling was performed.

Opsomming (max. 500 words)

Malaria is 'n peslike siekte wat geassosieer word met beduinde sosio-ekonomiese implikasies vir sub-Sahara Afrika. Die siekte word veroorsaak deur 'n bloed infeksie van die enkel sellulêre parasiet van die *Plasmodium* genus. Twee ensieme, naamlik Spermidien Sintetase en Proteïen Farnesieltransferase, is geïdentifiseer om noodsaaklik te wees vir die oorlewing van die parasiet. Die doelwit van hierdie verhandeling is die soektog en sintese van oorspronklike inhibeerders van hierdie twee ensieme met 'n sterk fokus daarop om struktuur/aktiwiteit interaksies te verstaan.

Om die eerste doelwit te bereik is molekulêre modellering toegepas. 'n Indiepte ondersoek word voorgestel om die onderliggende beginsels relevant tot hierdie tak van berekenkundige chemie te beskryf. Dit verseker dat eksperimente wat op hierdie tegnieke berus korrek opgestel word en dat die resultate binne konteks geïnterpreteer word. Twee virtuele hoë-deurset skerms was deurgevoer op voorbereide kristallografiese strukture van Spermidien Sintetase. Die eerste het berus op 'n farmakoforiese metode en die tweede op LibDock. 'n Self-ontwikkelde sagteware gereedskap stuk is gebruik om a databasis van 7.1 miljoen verbindings te filtreer voor dit gebruik is in hoë-deurset skerms. Uiteindelik is CDOCKER gebruik om die potensiele aktiwiteit van "treffer" verbindings te beraam.

Spermidien syntetase het 'n natuurlike affiniteit vir adenosien en hierdie eienskap is benut deur analoeë af te lei na potensiele inhibeerders teen die ensiem. Dit is bewerkstellig deur die insluiting van beide elektrofiliese asook nukleofiele funksionele groepe op gekose posisies. Dit het die gebruik van 'n hoë opbrengs Mitsunobu reaksie ingesluit. 'n Aantal ander addisionele residueë is toe gesintetiseer en geheg aan die afgeleide adenosien om die ensiem setel te vul en sodoende die affiniteit te verhoog.

Vir die tweede ensiem wat geteiken is vir inhibisie, Proteïen Farnesieltransferase, is indool benuttig as 'n begin steier te dien om potensiele treffers *de novo* te sintetiseer. Dit is geteiken om die indool af te lei op die N en 3' posisies en die kristal struktuur van een van hierdie tussengangers is gepubliseer. Verder is 'n sintetiese weg, wat uitloop het op 'n palladium gekataliseerde Suzuki koppeling, uitgevoer.

Acknowledgements

My family, whom I love without measure. For being there for me – always.

Dr Stephen Pelly, for guidance - both academic and otherwise

Prof Willem van Otterlo, for sharing his enthusiasm

The members of GOMOC (and Kaliefie) for sharing coffee

Dr Vincent Smith – For help in getting that ACTA out

Stellenbosch University– For being a world class institution

CHPC – For delivering world class software

TIA – For financial support

NRF – For financial support

Table of contents

1. Introduction	1
Medicinal chemistry	1
The drug discovery process	1
Targets.....	2
Hit discovery strategies	4
Rational drug design.....	4
Virtual High Throughput Screening	6
High throughput screening.....	6
Selective optimisation of side affects	7
“Me too” drugs.....	7
Hit to lead.....	7
Pharmacodynamics	7
Pharmacokinetics - ADMET	9
Bioisosterism.....	14
Clinical trials and moving from bench to bedside	15
2. Malaria.....	17
Introduction and life cycle.....	17
Current antimalarials and their mechanisms of action	18
4-Aminoquinolines	18
8-Aminoquinolines	19
Arylaminoalcohols.....	19
Artemisinin	20
Antifolates	21
Respiratory chain inhibitors.....	21
Antibiotics	22
Inhibition targets for this dissertation.....	23
Polyamine biosynthesis - Spermidine synthase.....	23
Protein prenylation - Farnesyltransferase	26
Aims.....	27
3. Molecular modelling – General.....	28
Molecular mechanics	28
Forcefields and CHARMM.....	28
Solvent models	31
No solvation model	31
Implicit solvent models	31
Explicit solvent model.....	32
Minimisation	33
High temperature molecular dynamics	34
Monte Carlo simulated annealing	34
Docking.....	34
Virtual high throughput screening and scoring	35
Databases.....	35
Docking Algorithms	36
Scoring.....	37
4. Molecular Modelling - Spermidine synthase.....	39
2PT9.....	39
Protein preparation	39
Virtual high-throughput screens.....	43
217C	47
Protein preparation	47
Hotspot guided screen - LibDock.....	48
Database preparation tool	48
Scoring functions.....	51
Libdock sequence	53
5. Spermidine Synthase – Electrophilic adenosine.....	56
Design.....	56

Synthesis – Chloroadenosine.....	58
Synthesis – RHS moieties.....	62
Reduction.....	62
Chemoselective protection.....	62
Aldehyde to thiol conversion.....	63
Synthesis – Linking attempts1.....	66
Conclusion.....	68
6. Spermidine Synthase – Nucleophilic adenosine.....	69
Synthesis – Thioester adenosine.....	69
Synthesis – RHS moieties.....	71
Method development using alkyl chains.....	71
AdoDATO and 4MCHA hybrid.....	75
Tyrosine derivative.....	79
Triazole derivatives.....	83
Results.....	90
Conclusion.....	91
7. Protein Farnesyltransferase inhibition.....	92
Derivatisation of indole <i>N'</i> position – Proof of concept.....	92
Synthetic planning.....	94
Synthesis.....	95
An alternative strategy.....	97
Derivatisation indole 3' position – (Xylene).....	100
Familiarisation of the biochemical environment.....	100
Planning.....	103
Synthesis.....	104
Derivatisation indole 3' position – (Maleimide).....	107
Planning.....	107
Synthesis.....	108
Conclusion.....	110
8. Experimental.....	111
Synthetic Procedures - General.....	111
Solvent purification.....	111
Chromatography.....	111
Spectroscopy.....	111
Spectrometry.....	111
General.....	111
Synthetic Procedures (Chapter 5).....	112
3 ((3 <i>aR</i> ,4 <i>R</i> ,6 <i>R</i> ,6 <i>aR</i>)-6-(6-Amino-9 <i>H</i> -purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)methanol.....	112
3 (Alternative synthesis).....	112
4 9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-6-(Chloromethyl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine.....	113
6 (4-Chlorophenyl)methanol.....	113
8 <i>Tert</i> -butyl (2-hydroxyethyl)carbamate.....	114
9 2-((<i>Tert</i> -butyldimethylsilyl)oxy)ethanamine.....	114
11 4-((<i>Tert</i> -butyldimethylsilyl)oxy)benzaldehyde.....	115
12 (4-((<i>Tert</i> -butyldimethylsilyl)oxy)phenyl)methanol.....	115
13 <i>S</i> -4-(<i>Tert</i> -butyldimethylsilyl)benzyl ethanethioate.....	115
14 (4-((<i>Tert</i> -butyldimethylsilyl)oxy)phenyl)methanethiol.....	116
Synthetic Procedures (Chapter 6).....	117
20 <i>S</i> -(((3 <i>aS</i> ,4 <i>S</i> ,6 <i>R</i> ,6 <i>aR</i>)-6-(6-Amino-9 <i>H</i> -purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)methyl)ethanethioate.....	117
21 9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-2,2-Dimethyl-6-((octylthio)methyl)tetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine.....	118
22 (2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i>)-2-(6-Amino-9 <i>H</i> -purin-9-yl)-5-((octylthio)methyl)tetrahydrofuran-3,4-diol.....	119
23 9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-6-((Hexylthio)methyl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine.....	119
24 (2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i>)-2-(6-Amino-9 <i>H</i> -purin-9-yl)-5-((hexylthio)methyl)tetrahydrofuran-3,4-diol.....	120
25 9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-2,2-Dimethyl-6-((octan-2-ylthio)methyl)tetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine.....	121
26 (2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i>)-2-(6-Amino-9 <i>H</i> -purin-9-yl)-5-((octan-2-ylthio)methyl) tetrahydrofuran-3,4-diol.....	121
27 9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-6-((Heptan-2-ylthio)methyl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine.....	122
28 (2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i>)-2-(6-Amino-9 <i>H</i> -purin-9-yl)-5-((heptan-2-ylthio)methyl) tetrahydrofuran-3,4-diol.....	122
33 4-(4-Nitrophenyl)butyl methanesulfonate.....	123

32	9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-2,2-Dimethyl-6-(((4-(4-nitrophenyl)butyl)thio)methyl) tetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine	123
39	(<i>S</i>)-Methyl 2-amino-3-(4-((<i>tert</i> -butyldimethylsilyloxy)phenyl)propanoate	124
40	(<i>S</i>)-Methyl 2-((<i>tert</i> -butoxycarbonyl)amino)-3-(4-((<i>tert</i> -butyldimethylsilyloxy)phenyl)propanoate	124
44	Propane-1,3-diyl dimethanesulfonate	124
45	3-Azidopropyl methanesulfonate	125
46	9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-6-(((3-Azidopropyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine	125
47	9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-6-(((3-(4-(Aminomethyl)-1 <i>H</i> -1,2,3-triazol-1-yl)propyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine	126
48	(2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i>)-2-(6-Amino-9 <i>H</i> -purin-9-yl)-5-(((3-(4-(aminomethyl)-1 <i>H</i> -1,2,3-triazol-1-yl)propyl)thio)methyl)tetrahydrofuran-3,4-diol	127
50	Ethane-1,2-diyl dimethanesulfonate	127
51	2-Azidoethyl methanesulfonate	127
52	9-((3 <i>aR</i> ,4 <i>R</i> ,6 <i>S</i> ,6 <i>aS</i>)-6-(((2-Azidoethyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4- <i>d</i>][1,3]dioxol-4-yl)-9 <i>H</i> -purin-6-amine	128
53	(1-(2-(((3 <i>aS</i> ,4 <i>S</i> ,6 <i>R</i> ,6 <i>aR</i>)-6-(6-Amino-9 <i>H</i> -purin-9-yl)-2,2-dimethyltetrahydrofuro [3,4- <i>d</i>][1,3]dioxol-4-yl)methyl)thio)ethyl)-1 <i>H</i> -1,2,3-triazol-4-yl)methanaminium	129
54	(2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i>)-2-(6-Amino-9 <i>H</i> -purin-9-yl)-5-(((2-(4-(aminomethyl)-1 <i>H</i> -1,2,3-triazol-1-yl)ethyl)thio)methyl)tetrahydrofuran-3,4-diol	130
	Synthetic Procedures (Chapter 7)	131
58	<i>Tert</i> -butyl 4-(hydroxymethyl)-1 <i>H</i> -imidazole-1-carboxylate	131
57	<i>Tert</i> -butyl 4-(chloromethyl)-1 <i>H</i> -imidazole-1-carboxylate	131
61	<i>Tert</i> -butyl 5-cyano-1 <i>H</i> -indene-1-carboxylate	132
64	(1-Methyl-1 <i>H</i> -imidazol-5-yl)methanol	132
65	5-(Chloromethyl)-1-methyl-1 <i>H</i> -imidazole hydrochloride	133
62	1-((1-Methyl-1 <i>H</i> -imidazol-5-yl)methyl)-1 <i>H</i> -indole-5-carbonitrile	133
64	<i>Tert</i> -butyl 5-bromo-1 <i>H</i> -indole-1-carboxylate	133
65	Acetoxy(5-bromo-1-(<i>tert</i> -butoxycarbonyl)-1 <i>H</i> -indol-3-yl)mercury	134
66	(5-Bromo-1-(<i>tert</i> -butoxycarbonyl)-1 <i>H</i> -indol-3-yl)boronic acid	134
71	Tris(dibenzylideneacetone)dipalladium	135
72	5-Bromo-1-tosyl-1 <i>H</i> -indole	135
79	5-Bromo-1 <i>H</i> -indole-3-carbaldehyde	136
78	5-Bromo-3-(2-nitrovinyl)-1 <i>H</i> -indole	136
80	<i>Tert</i> -butyl 5-bromo-3-formyl-1 <i>H</i> -indole-1-carboxylate	137
	Modelling SpdSyn – 2PT9	138
	Protein preparation	138
	vHTS – Pharmacophore method	140
	Modelling SpdSyn – 2I7C	141
	Protein preparation	141
	Hotspot guided screen - LibDock	142
	CDocking of 1600 LibDock results	143
	SpdSyn docking parameters for designed compounds	143
	Modelling PFT – 2R2L	144
	PFT protein preparation	144
	2R2L docking protocols	145
9.	References	146

List of Figures

Figure 1-1	Pie chart of the main biomolecule classes which are targeted by current pharmaceuticals. ¹²	3
Figure 1-2	Three of the statins	4
Figure 1-3	Saquinavir	5
Figure 1-4	Graph of the total number of PDB entries per year	6
Figure 1-5	Dose response curves highlighting differences between efficacy and potency.	8
Figure 1-6	Examples of typical Phase II metabolism conjugation reactions	12
Figure 1-7	Graphical representation of the therapeutic window.	13
Figure 1-8	Histogram indicating how common it is for drugs to fail in the clinical trials. ¹⁰	15
Figure 2-1	Lifecycle of Plasmodia species. ⁶⁷	18
Figure 2-2	4-Aminoquinolines	19
Figure 2-3	8-Aminoquinolines	19
Figure 2-4	Arylaminoalcohols	20
Figure 2-5	Artemisinin and derivatives	21
Figure 2-6	Antifolates	21
Figure 2-7	Respiratory chain inhibitors	21
Figure 2-8	Antibiotics	22
Figure 2-9	Active site of SpdSyn	25
Figure 3-1	Simple depiction of how various atoms are assigned parameters ¹⁵¹	29
Figure 3-2	Graphical plot illustrating the Lennard Jones potential. ¹⁵³	31
Figure 3-3	Simplified depiction of the energy landscape as function of a single variable.	33
Figure 3-4	CDOCKER protocol workflow ¹⁶⁶	35
Figure 3-5	Example of hotspots as generated and used by LibDock ¹⁷¹	37
Figure 4-1	4MCHA and AdoDATO	39
Figure 4-2	Optimisation required for hydrogen atoms	40
Figure 4-3	Increasing the total energy of the system by the increase of the temperature.	41
Figure 4-4	Kinetic equilibration of the introduced energy.	41
Figure 4-5	Graph of the potential energy as the production run of the standard dynamics cascade proceeds.	42
Figure 4-6	Example of histidine charge / tautomeric optimisation	42
Figure 4-7	Successful redock of 4MCHA (in orange) back into 2PT9.	43
Figure 4-8	Development of the pharmacophore inspired by 4MCHA (carbons in orange)	45
Figure 4-9	Completed pharmacophore inspired by 4MCHA (carbons in orange).	45
Figure 4-10	Prepared crystal structure of 2I7C	48
Figure 4-11	Schematic of Pipeline Pilot workflow to prepare large databases for screening.	49
Figure 4-12	Different scoring functions applied to the various potential inhibitors reported by Jacobsson <i>et al.</i> ¹⁸⁵	52
Figure 5-1	Natural substrates of SpdSyn	56
Figure 5-2	AdoDATO interactions areas with SpdSyn	56
Figure 6-1	First iteration of a hybrid molecule	76
Figure 6-2	Ideal hybrid 30	76
Figure 6-3	Proposed hybrid of AdoDATO and 4MCHA docked into the active site of SpdSyn.	77
Figure 6-4	Proposed tyrosine hybrid derivative 36 and tyrosine	79
Figure 6-5	Proposed hybrid of 4MCHA, 36, docked into the active site of SpdSyn.	80
Figure 6-6	The 3 carbon linked triazole derivative, docked into the active site	83
Figure 6-7	Two possible triazole regioisomers	87
Figure 6-8	SpdSyn inhibition results	90
Figure 7-1	Tetrahydroquinoline based inhibitor developed by van Voorhis <i>et al.</i> and our indole derivative. ¹⁴⁵	93
Figure 7-2	Wide entrance to the active site of protein farnesyltransferase (PDB code: 2R2L).	93
Figure 7-3	NMR segment of the <i>tert</i> -butyl group	96
Figure 7-4	Alternative proof of concept compound	97
Figure 7-5	Crystal structure of the proof of concept compound, 62.	99
Figure 7-6	Superimposition of the active site of 18 PFT structures, co-crystallised with various ligands.	101
Figure 7-7	Demonstration of possible rotamer creation	102
Figure 7-8	Active site surface of 2R2L docked with the planned indole-xylene derivative 63.	102
Figure 7-9	Mechanism of indole 3' position's nucleophilicity	103
Figure 7-10	Maleimide-indole 73 docked in the active site of PFT	107

List of Tables

Table 2-1	Important crystal structures of SpdSyn.....	25
Table 2-2	Notable crystal structures of mammalian PFT.....	27
Table 3-1	Molecular databases.....	36
Table 3-2	Scoring functions available in DS.....	38
Table 4-1	Results of the Sigma-Aldrich catalogue screen by pharmacophoric features.....	46
Table 4-2	SMARTS filters.....	50
Table 4-3	Database preparation parameters.....	51
Table 4-4	Range of magnitudes of the scores which delivered potent results.....	52
Table 4-5	Possible hits of the vHTS using LibDock from 3.2 million compounds from ChemDB database.....	54
Table 5-1	¹ H NMR spectroscopic signals of 3 and 4.....	61
Table 6-1	Assignment of 20 ¹ H & ¹³ C NMR spectroscopy signals.....	70
Table 7-1	Energy differences in crystal structure and the global minimum.....	100

Glossary

Hit.....	A primary, active compound, with non-promiscuous binding behaviour, exceeding a certain threshold value in a given assay(s) ¹
Combinatorial Chemistry.....	Synthesis technologies to generate compound libraries rather than single products ¹
Conformer	A molecule in a specific conformation
DtBAD	di- <i>tert</i> -butyl azodicarboxylate
Drug-likeness	A computational scoring metric for the similarity of a given structure to a representative reference set of marketed drugs.
Lead	Prototypical chemical structure or series of structures that demonstrate activity and selectivity in a pharmacological or biochemically relevant screen ¹
Metabolite	a chemical compound which has been (partially) metabolised
Peroral	performed or administered through the mouth
Pharmacodynamics.....	The study of how ligands interact with their target binding site
Pharmacokinetics.....	The study of drug absorption, drug distribution, drug metabolism and drug excretion
Pharmacophore.....	Ensemble of steric and electronic features that is necessary to ensure optimal interactions with a specific biological target structure and to trigger (or to block) its biological response
Pose.....	A molecule which is in a specific conformation and position relative to a ligand
Prodrug	A drug which is administered in a non-active form to become metabolised to the active drug
Soft drug	A drug purposefully designed to have non-toxic metabolites

Abbreviations

ACT	Artemisinin combination therapy
AdoDATO	S-adenosyl-1,8,diamino-3-thioactane
AdoMet	S-adenosylmethionine
AdoMetDC	S-adenosylmethionine decarboxylase
Boc	<i>tert</i> -butyl carbamate
Boc ₂ O	di- <i>tert</i> -butyl dicarbonate
CDOCKER	CHARMm-based DOCKER
CFF	Consistent Force Field
CG	conjugate gradient (minimisation algorithm)
dcAdoMet	decarboxylated S-adenosylmethionine
DCM	dichloromethane
DDDE	distance dependant di-electrics (implicit solvent method)
DMAP	(dimethyl)aminopyridine
DMP	dimethoxy propane
DS	Accelrys Discovery Studio 3.1 and 3.5
DiBAD	di- <i>tert</i> -butyl azodicarbonate
eIF5	eukaryotic initiation factor 5
FPPIX	ferriprotoporphyin IX
GB	Generalised Born
GGI	geranylgeranyl transferase I
GGII	geranylgeranyl transferase II
GPCR	G-protein coupled receptor
HMPA	Hexamethylphosphoramide
LAH	lithium aluminium hydride (LiAlH ₄)
4MCHA	4-methylcyclohexylamine
MMFF	Merck molecular force field
MTA	5'-methylthioadenosine
NME	new molecular entity
NMR	nuclear magnetic resonance
ODC	ornithine decarboxylase
PAO	polyamine oxidase
PBSA	Poisson-Boltzmann with non-polar surface area
PDB	protein databank
pfAdoMetDC/ODC	<i>Plasmodium falciparum</i> adenosine methionine decarboxylase / ornithine decarboxylase
PME	Particle Mesh Ewald
PFT	Protein Farnesyltransferase
PTSA	p-toluene sulfonic acid
Putrescine	1,4-diaminobutane
SAR	structure activity relationship
SpdSyn	spermidine synthase
hsSpdSyn	<i>Homo sapiens</i> spermidine synthase
pfSpdSyn	<i>Plasmodium falciparum</i> spermidine synthase
QSAR	Quantitative structure-activity relationship
SBDD	Structure based drug design
SD	steepest descent (minimisation algorithm)
SpmSyn	spermine synthase
TBDMS(Cl)	<i>tert</i> -butyl dimethyl silyl(chloride)
TEA	triethylamine
THF	tetrahydrofuran
THQ	tetrahydroquinoline
TLC	thin layer chromatography
vHTS	virtual high throughput screen

Conventions used throughout this dissertation

3D Structures are presented as follows (unless specified otherwise):

Atoms

Carbon: Grey (or orange for highlighting).

Oxygen: Red

Nitrogen: Blue

Hydrogen: White

Sulfur: Yellow

Protein structures

Alpha-helices..... Red and

Beta-sheets Light blue.

Other

Hydrogen bonding..... Dotted green lines

Pi-stacking Orange lines

1. Introduction

Medicinal chemistry

The use of medicine has been around since records began. In ancient Egypt, people consumed extracts of the poppy plant to alleviate pain, not knowing how or why they worked. This extract is still widely used today, which is more commonly known as morphine.² In ancient Greece, women dilated their pupils to beautify themselves with *Solanaceae* plants which contained atropine.³ Aztec and Mayan civilisations used the ipecac root, which contained emetine, to treat infections and induce vomiting.⁴ The extract of the Willow tree's bark, rich in salicylic acid, was described by Hippocrates to reduce pain and fever.⁵ In 1928 Alexander Fleming made the Nobel Prize winning discovery of Penicillium and so ushered humanity into the antibiotic era.^{6,7}

Throughout the twentieth century, our brightest minds have modified natural product isolates to create life altering pharmaceuticals. Although we still look towards Nature for inspiration, newer methods are being developed. Today, modern drug discovery programmes rely on *ab initio* principles to rationally design pharmaceuticals. This is the field of modern medicinal chemistry.

The drug discovery process

Modern drug discovery is an extended, complicated and extremely expensive process – the average cost to bring a drug to market after a hit compound has been identified is estimated to be \$1.8 Billion.⁸ However, the profit incentives are equally lucrative, in addition there is satisfaction of being involved in a project which may benefit the lives of millions of people. To bring the entire project to fruition, the skills of multiple disciplines are required – those of science, medicine, law and business economics.

The first step in the endeavour is the selection of a disease. Different diseases present different difficulties, incentives, success rates and capital requirements (both financial and human). For example, diseases of the third world affect many millions of people who can barely afford food – let alone expensive pharmaceuticals. This demands that the cost for each therapeutic administration be very inexpensive. In this case, the main incentive would be the humanitarian effort. However, drugs which alleviate ailments of the first world can command wide profit margins. Pills which offer weight loss are marketed towards patients who can typically pay for expensive, tailor-made pharmaceuticals. In this case, the main incentive is profit.

Key to the industry are patents: Patent protection only lasts 20 years whereas drug development typically lasts roughly 13 years – meaning there are only 7 years to recoup the massive expenditures and still turn a profit.⁸ Sometimes, development takes longer and the project may be cancelled. In light of this, countries (including USA) have realised that this is stifling the output of new drugs and now offer market exclusivity periods (5 years for a New Chemical Entity) to incentivise new innovators.⁹

After a disease has been selected, the biological target for manipulation needs to be decided upon. Following this, a number of compounds are tested for activity. Once a compound exhibiting a measure of desired activity is discovered, known as a “hit”, it becomes the subject of intense further development. This hit is then developed in the “hit to lead” phase: This involves significant synthetic work which is performed in parallel with animal studies to optimise binding characteristics and assess toxicity. If these succeed, human clinical trials may commence. The clinical trials are notoriously rigorous and most candidates fail them.¹⁰ However, should the drug candidate succeed the clinical trials, the drug is registered and placed on the market.

Concurrent with the trials, large scale industrial syntheses of the drugs are developed. During this time, costs soar. Strategic business decisions need to be made; shareholder equity is diluted to ensure that liquid capital is always on tap. This while keeping in mind that at any time, the drug candidate may fail a step. At best, the process may be moved back several steps. At worst, the entire project may be terminated and all investments made are forfeit.¹¹ Indeed, the road to launch a new drug is fraught with significant hurdles which require careful consideration. One of the first ones is the selection of an appropriate target.

Targets

In order for a pharmaceutical to exert its action, it interacts with the naturally occurring physiological effectors. These effectors, deemed drug targets, are chosen for their key involvements in cellular processes. The major targets for pharmaceuticals and their prevalence are presented in Figure 1-1. The largest group of targets are receptors.¹² These proteins are key components in signal transduction. They are embedded within the cellular membrane where they are stimulated by extracellular, chemical messenger molecules, *hormones*, and relay these messages into the cell. Within this class, the G-protein coupled receptors (GPCR's) form the largest subset. Other receptors include the enzyme linked tyrosine kinases and ion-channels.¹³⁻¹⁶ The next biggest target for pharmaceuticals are enzymes. Enzymes are the main effectors of regulating the cellular environment, being intricately involved in the conversion of metabolites. Other targets include nucleic acids or the factors which govern their transcription.

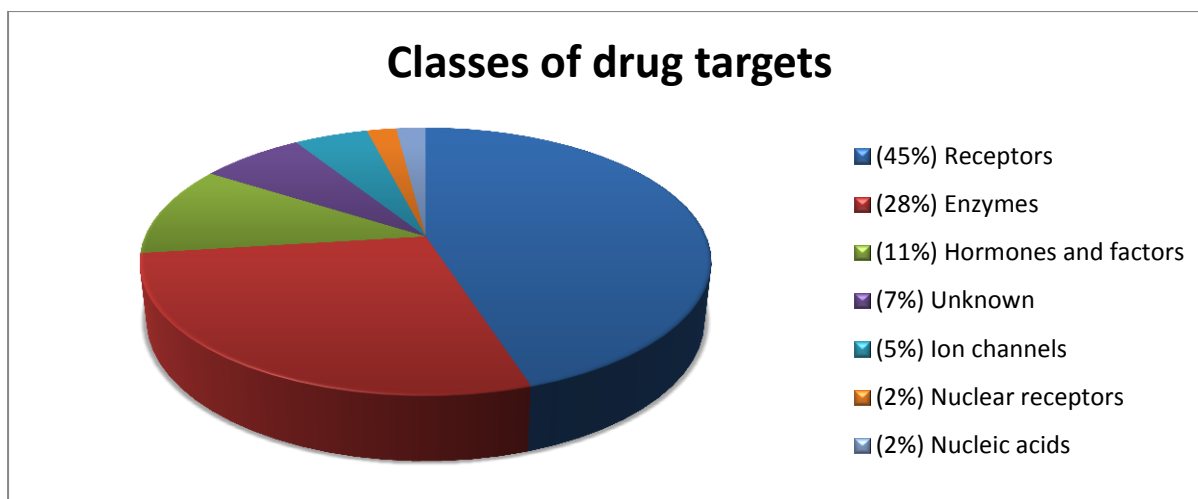


Figure 1-1 Pie chart of the main biomolecule classes which are targeted by current pharmaceuticals.¹²

Although there are exceptions, there are 2 major strategies to modulate the cellular environment for a wanted physiological response: Either through the decrease or increase in the activity of these mentioned targets.¹⁷ Compounds which have these effects on targets are termed antagonists and agonists, respectively. Typically, this is achieved through the use of molecules which resemble the natural substrates which are termed substrate analogues.¹⁸

In the former case, inhibitors are effectively used to occupy the active site. This blocks the natural substrate from entering and so limits product formation. Take for example, the sulfonamides: These antibiotics inhibit the enzyme dihydropteroate synthetase, stopping folate biosynthesis. The effect is that the bacterium becomes starved of essential folates, dies and so the infection is ended. Because humans do not have this enzyme and folates are acquired through the diet, only the bacterium is affected.¹⁹ Inhibition can be classed into three major categories. Firstly, there is competitive inhibition. This is when the drug competes with the natural substrate for the active site of the target. Secondly, uncompetitive inhibition requires that the natural substrate be bound prior to the drug's interaction. Thirdly, there is the case of irreversible inhibition. In this case, a covalent bond is formed between the drug and enzyme, which permanently disables the enzyme. Typically, nucleophilic residues (serine and cysteine, with hydroxyls and thiols, respectively) in the active site perform nucleophilic attack on the drug's electrophilic centres. Penicillin and aspirin are well known examples which fall within this category.^{20, 21}

In contrast to inhibiting targets, it is also possible to stimulate them. Agonists can exert their effects either by direct or allosteric interaction with the active site.

When deciding upon a target to treat a particular disease, careful consideration should be used. Altering a biological pathway or signal cascade may incur numerous unwanted side effects and the implications thereof must be fully understood. For example, in the design of cholesterol medication, an unforeseen problem arose. Although cholesterol is supplemented from diet, it is also synthesised by the body. A number of enzymes are involved in the biosynthetic pathway and any one of them can be subjected to inhibition. One effort to block this cholesterol producing pathway was to inhibit one of the final enzymes in the pathway. Although the enzyme was successfully inhibited and cholesterol production within the body was significantly lowered, toxicity issues were reported. It was soon determined that the substrate to the enzyme that was being inhibited was, similar to cholesterol itself, a very hydrophobic molecule. This led to the accumulation of the insoluble intermediates which, in turn, led to the reported toxicity issues. By choosing an earlier step in this pathway, the accumulating intermediates were much more water soluble and did not exhibit the toxicity issues. Ultimately, this second strategy led to the statins, one of the most profitable class of pharmaceuticals on the market, some of which can be seen in Figure 1-2.¹⁸

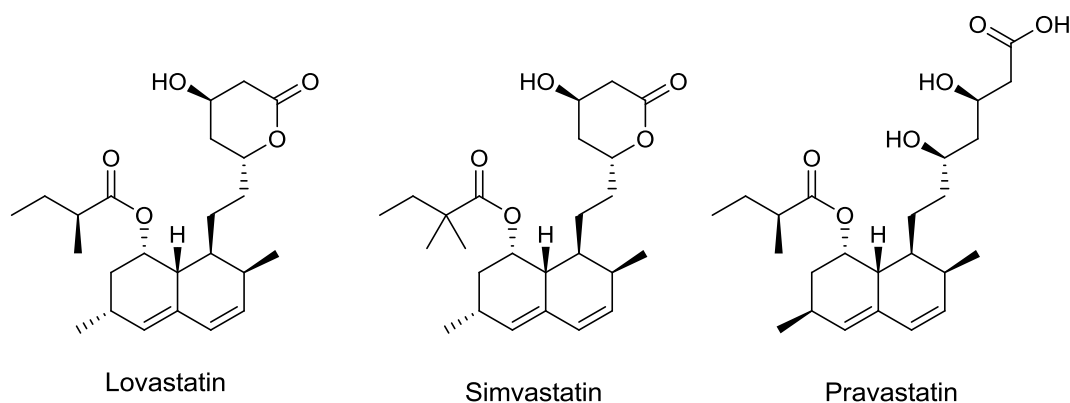


Figure 1-2 Three of the statins

Hit discovery strategies

After having selected a target in the drug discovery pipeline, the next step entails finding a molecule which has some sort of modulating activity against the target. This requires the discovery of a *hit*. A hit is a compound which shows notable activity towards a particular target. To obtain a hit, there are a number of strategies.

Rational drug design

Rational drug design entails tailoring a compound to interact with a specific target. This can either be accomplished by *ligand* based drug design (LBDD) or *receptor* based drug design, more commonly known as structure based drug design (SBDD). The former case requires a set of active ligands. These ligands are systematically modified and their activities measured

to build a map of the functional chemical space around the target. By compiling this information into a statistical model, it is possible to predict which compounds would also be active. This is known as a quantitative structure activity relationship (QSAR) strategy.²² The successful QSAR application relies on the two parameters regarding the input training set. The training set should be large, preferably above 100 compounds. Secondly, the set should include ligands which are as structurally diverse as possible. Failure of incorporating diverse ligands would lead to diminished possibility of finding a novel structure with a significant increase in pharmacological activity.

On the other hand, SBDD requires the 3D co-ordinates of the target protein, obtained by X-ray crystallography or NMR spectroscopy. This strategy is most likely to yield novel drug classes as the structural features of the target site can be thoroughly and accurately explored. A recent success using this method is exemplified in the marketed anti-HIV pharmaceuticals Saquinavir, Ritonavir, Indinavir and Nelfinavir (see Figure 1-3).²³ These inhibitors target the retroviral protease of the human immunodeficiency (HIV) virus type 1. Using structures which were obtained from X-ray crystallography, it was observed that the HIV protease cleaves an amide bond between Tyr-Pro and Phe-Pro peptide sequences. The designers then incorporated a structural motif which has a similar interaction to the protein, without proteolytic susceptibility. This then led to a high degree of affinity for the key enzyme and significantly reduced viral reproductive capability.

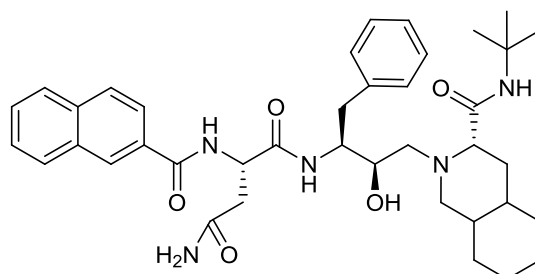


Figure 1-3 Saquinavir

If a structure is unavailable, a homology model can be constructed. A homology model is a protein model which is constructed piecewise from other known protein segments. Although there is less certainty in these models, compared to that of an experimental observation, these models have successfully been used in the design of pharmaceuticals.¹⁸

One of the most significant methods available in SBDD is the idea of fragment growing. Small molecule fragments are placed within the active site. These fragments are then linked together to grow a complete molecule. This method has the advantage that new scaffolds are more likely to be introduced. Programs which do this include CAVEAT, SPROUT and LUDI.²⁴⁻²⁶

In light of the importance which knowledge of the target structure may bring, the protein databank (PDB) was founded.²⁷ The PDB is a global database of protein structures which have been determined either through crystallography or NMR. Since it was founded, the PDB has grown tremendously – a testimony to the importance of its existence (see Figure 1-4). Further, 92% of marketed drugs have targets which are similar to proteins found within the PDB.²⁸

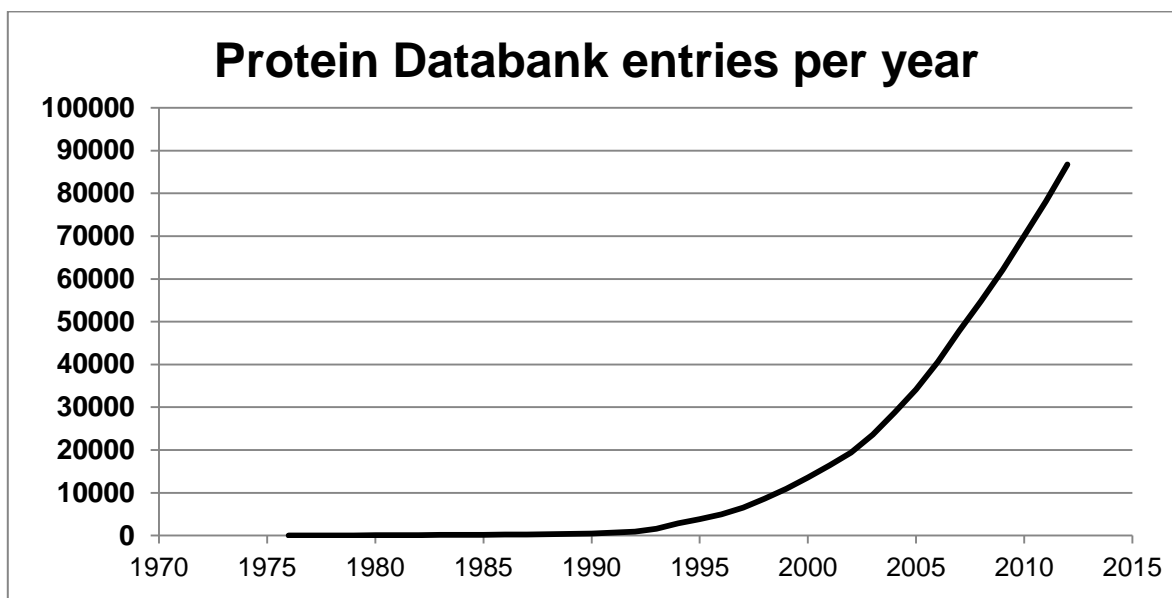


Figure 1-4 Graph of the total number of PDB entries per year.

Virtual High Throughput Screening

The other main hit discovery strategy which uses the 3D structure of a protein *in silico*, is virtual high throughput screening (vHTS). In contrast to the piecewise construction of new molecules *de novo*, vHTS attempts to place whole molecules in the active site and then use scoring algorithms to measure how well they fit.^{29, 30} Using pre-compiled databases, many millions of compounds may be screened for possible activity.³¹⁻³³ This topic is discussed in more detail in chapter 4.

High throughput screening

High throughput screening (HTS) entails physically testing large compound libraries for possible bioactive qualities *in vitro*. These tests are performed by autonomous robots which are able to screen over a thousand micro assays on a single multiwell plate. Aside from the robotic requirements, there are two main caveats of this technology: The availability of a suitable bioassay and the libraries themselves.

Firstly, the assay should lend itself to high throughput implementation. Typically, UV/Vis spectroscopy and radioactivity measurements are employed. Secondly, the libraries should

be both sufficiently large (>10 000 compounds) and diverse to thoroughly explore the chemical space of the target. In industry, these libraries may comprise previously synthesised molecules or be obtained through synthesis.¹⁸

Selective optimisation of side effects

Selective optimisation of side effects (SOSA) attempts to capitalise on an unexpected but useful side effect in one drug making it the primary in another. Ideally, the original activity of the drug is minimised by optimisation of the drug molecule.³⁴ The drug sildenafil, more popularly known as Viagra®, was developed by this method.

“Me too” drugs

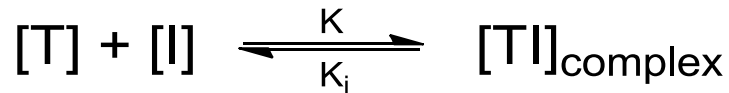
By starting with an already marketed drug, a competing company may attempt producing its own analogue of similar structure. The newly developed drug may offer improvements over the current market leader. It is interesting to note that between 1989 and 2000, 76% of new molecular entities (NME) were targeting already targeted protein domains, compared to only 6% which targeted new protein domains.²⁸ Sir James Black, a Nobel laureate in medicine, once said: “The most fruitful basis for the discovery of a new drug is to start with an old drug.”³⁵

Hit to lead

Once a hit has been discovered, it becomes the subject of intense development. A number of properties need to be optimised. First and foremost, the lead candidate needs to have maximum affinity to the target and this needs to be developed. The study of how the drug interacts with a target site is defined as pharmacodynamics. But also, having successfully developed a compound with maximal activity *in vitro* is most definitely not a final pharmaceutical product. Delivering a chemical compound into someone’s body is fraught with complexity. The body also affects the drug in numerous ways. As is often the case, it is not the most potent compound which is placed on the market, but a compound which also satisfies numerous other criteria. These criteria are categorised under *pharmacokinetics*. Firstly, we shall discuss pharmacodynamics.

Pharmacodynamics

Pharmacodynamics concerns the aspects which allow a drug to bind to its target. Before we can discuss drug binding, we need a way to quantify the interaction. The majority of drugs have reversible interactions with their targets. The equilibrium between the enzyme and the inhibitor can be depicted as follows:



Where [T] & [I] are the free target & inhibitor concentrations; [TI] is the concentration of the complex formed; and K & K_i are the forward & reverse equilibrium constants, respectively. As the K_i decreases, the equilibrium shifts to the right, favouring the inhibitor complex – indicating increased binding affinity. The average K_i of marketed pharmaceuticals are in the order of 20 nM.²⁸ Alternatively, an inhibitor's potency can be quantified by an IC_{50} – the concentration where 50% of maximal inhibition is observed. Graphically, this may be visualised by setting up a dose response curve (see Figure 1-5).

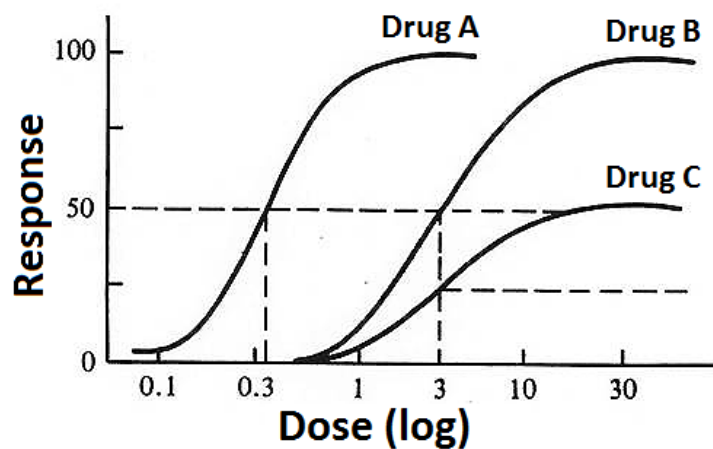


Figure 1-5 Dose response curves highlighting differences between efficacy and potency. Efficacy refers to the magnitude of the response obtained, regardless of drug concentration – in this case drugs A and B have similar efficacies. Potency refers to the concentration where half the maximal response is reached – in this case drugs B and C have similar potencies.

Typically, IC_{50} is used in whole cell assays, whereas K_i is more typically used in enzymatic assays. Through the use of the Cheng-Prusoff equation³⁶, the two may be interconverted.

$$K_i = \frac{IC_{50}}{\left(1 + \frac{[L]}{K_D}\right)} \quad \text{Eq 1}$$

Where [L] is the radio-ligand concentration used in the specific assay and K_D is the affinity for the radio-ligand.

Because K_i is the inverse of the equilibrium constant, it can be related to the change in Gibbs free energy by the following equation:

$$\begin{aligned} \Delta G &= -RT \ln K \\ &= RT \ln K_i \end{aligned} \quad \text{Eq 2}$$

Where ΔG equals the change in Gibbs free energy, R & T are the universal gas constants & temperature respectively, and K is the equilibrium constant.

The change in Gibbs free energy is more commonly given by the following expression

$$\Delta G = \Delta H - T\Delta S \quad \text{Eq 3}$$

Where ΔH and ΔS are the changes in enthalpy and entropy, respectively. In an attempt to quantify how various aspects of a drug binding to the target contribute to the change in free energy, the respective contributions are given by the following equation.³⁷

$$\Delta G = \Delta G_{\text{trans+rot}} + \Delta G_{\text{conf}} + \Delta G_{\text{polar}} + \Delta G_{\text{hydrophob}} + \Delta G_{\text{vdW}} \quad \text{Eq 4}$$

This allows us to optimise for each term individually when designing drug candidates. Keeping these parameters in mind, it is important to realise that drug binding is a relatively complex process:

Initially, the drug is in an entropically favoured, disorganised conformation as it tumbles freely in solution. The polar areas are complexed to solvent water molecules. For binding to be achieved, several things need to happen. Firstly, the molecule needs to adopt its bioactive conformation, which entails an entropic penalty. The drug then associates to the protein, losing its tumbling freedom – another entropic penalty. The desolvation penalty then comes into play: polar areas on the drug lose polar interactions as the water molecules around it are displaced. Even though these interactions may be replaced with chemical moieties from the protein, it is the *net energy difference* which affects the binding characteristics. Another favourable binding interaction then comes into play: the hydrophobic effect. Water molecules associated around hydrophobic areas of the drug are highly organised – an entropically disfavoured state. As the drug binds, these water molecules are then freed to resume a disorganised state. Finally, the van der Waals interactions govern the steric interactions as the drug complexes to the protein.³⁸

The interior of the receptor site may also be filled with water molecules. As the drug binds, these water molecules are displaced by the interactions of the protein with the drug.

Pharmacokinetics - ADMET

As a lead becomes further developed, it must always remain in the designer's mind that the ultimate goal is to administer the drug to a person. Patient safety is of paramount importance. To ensure this, the interactions a drug as it travels through the body must be studied and optimised in a predictive manner. These concerns are addressed under pharmacokinetics. Pharmacokinetics entails five criteria, collectively named ADMET: Absorption, Distribution, Metabolism, Excretion and Toxicity.

Absorption

For a drug to be effective, it needs to reach its site of action in its active form in sufficient quantity. But firstly, it needs to be administered. The preferred delivery method is oral, but it is also the most challenging route. The drug needs to be stable in the stomach acids, followed by the basic conditions of the gut. Following this, the drug needs to be hydrophobic enough to permeate the intestinal cellular membrane, yet still be hydrophilic enough to be transported in the blood. To better predict which compounds would show good absorptive qualities, Lipinski devised his rule of five.³⁹ This rule was obtained by studying the properties of numerous approved drugs which show good oral absorption. The rules state that ideally drugs should:

- have a molecular weight of less than 500;
- less than or equal 5 hydrogen bond donors;
- less than or equal 10 hydrogen acceptors;
- cLog P less than 5 (a measurement of hydrophobicity by the partitioning of the molecule between water and octanol, whereas the c refers to the computed estimate).

However, it should be noted that 20% of orally dosed drugs available on the market do not adhere to one of these parameters.²⁸ Supplementary to this work, Veber *et al.* suggested that the number of rotational bonds should not exceed 10 and that the total polar surface area should not exceed 140 \AA^2 .⁴⁰ To date the most accurate method to test *in vitro* whether lead candidates show good absorptive properties, is through the use of Caco-2 cell line test.⁴¹ Caco-2 is an immortalised line of intestinal cells which is able to effectively mimic *in vivo* conditions. After absorption from the intestines, the drug enters the hepatoportal vein to enter the liver. At this point, the drug endures the liver's so called "first pass effect" – to be subjected to liver metabolism before the drug may have reached its target. (Drugs which are injected, inhaled or taken transmucosally may bypass this effect.) Only then is the drug free to reach the heart and enter systemic circulation at physiological pH. By choosing alternate delivery methods, some of these difficulties may be avoided. These include: direct injection, topical application, transmucosal delivery and inhalation.

Should a drug be administered by direct injection, all of it would reach the blood stream and is, by definition, said to be 100% bioavailable. However, this may be drastically different for alternate administration routes. Some factors which might influence this include: the drug's physical properties, excipients used, physiological state and drug-drug interactions.⁴⁰

Distribution

For drugs to have any effect, they need to reach the target site and in sufficient quantity. By studying how the drug distributes around the body, it may be required to optimise the drug to target it to the correct locale. Drugs which struggle to cross the membrane may either be transported by macromolecules or be bound to blood albumin proteins. This reduces the amount of drug reaching the site of action. Drugs which are freely circulated in systemic blood are usually quickly removed by the kidneys or metabolised by the liver. However, if the drug is excessively hydrophobic, it may be absorbed and retained by fatty tissues. This may lead to sporadic and unplanned surges in drug levels. Crossing of the blood-brain barrier poses an additional challenge: there is an additional fatty layer to penetrate and this requires increased drug lipophilicity.^{42, 43} Similarly, crossing the placental barrier also has its own complications. Drugs which are present in maternal blood are easily transported to the foetal blood. However, the foetus does not have the same metabolic capabilities as the mother – with the implication that drugs can remain active for much longer. This demands special dosing requirements.^{44, 45}

Metabolism

Analogous to the immune system which defends the body against foreign biological entities (bacteria, proteins, etc.) the body is also equipped with a defence system against foreign small molecules, namely metabolism. Metabolism is the process by which foreign molecules, known as xenobiotics, are chemically altered so that they can safely be expelled from the body as metabolites. Even though pharmaceuticals may exhibit beneficial properties to the body, their metabolism is essential. Metabolic failure of a drug leads to its accumulation within the body and this, in turn, leads to toxicity issues. Conversely, if a drug is metabolised too quickly, it is expelled from the body before its beneficial function is bestowed.

The rate at which a xenobiotic is metabolised is described by its half-life - the period after which half of the initial concentration remains. This indicates the pharmaceutical's active duration. The fine tuning of a drug's half-life is an important step in the development process: Should a drug have an extended half-life, the drug would remain active for long periods of time. This is a wanted attribute for certain cases – for example behavioural changing pharmaceuticals which interact with the central nervous system. However, pain medication, which impairs lucidity, requires relatively quick half-life times to be able to better control the duration of this unwanted side effect.

Phase I metabolism of drugs is largely controlled by the cytochrome P450 isozymes. Although mainly located in the liver, they are found throughout the body. Although some these isozymes are involved in biosynthetic processes, they are also responsible for

xenobiotic metabolism.⁴⁶ This is achieved by the phase I metabolism of xenobiotics (oxidation, reduction and hydrolysis) to increase hydrophilicity for excretion via the renal system. The notable isozymes for human xenobiotic metabolism include: CYP3A4, CYP1A2, CYP2D6, CYP2E1 and CYP2C_x.⁴⁷ Even though the major pathways of metabolism can be predicted for these major groups, genetic polymorphism, induction & inhibition (affected by age,⁴⁸ ethnicity,⁴⁹ diet,⁵⁰ and health⁵¹) vastly increases the variability of how these enzymes interact with various xenobiotics.⁵²

In conjunction with phase I metabolism, phase II metabolism entails the conjugation of xenobiotics to hydrophilic moieties by various transferase enzymes. Typical examples of these transferases include: UDP-glucuronosyltransferase, Glutathione-S-transferase, *N*-acetyltransferase, Sulfotransferase (see Figure 1-6).⁵³

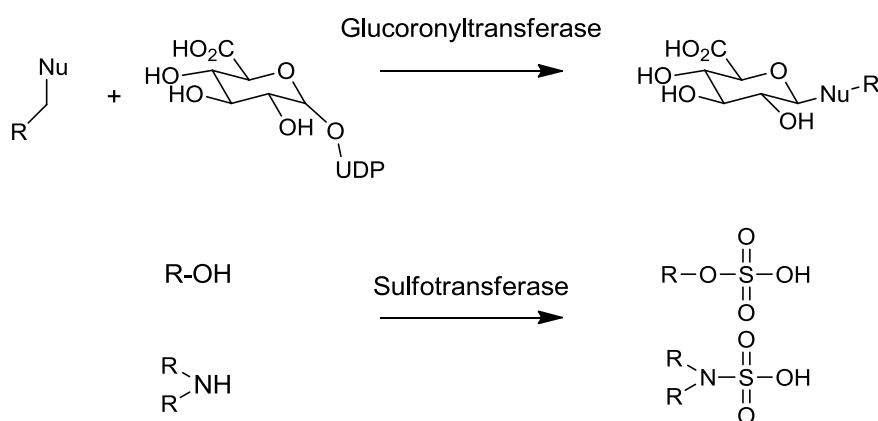


Figure 1-6 Examples of typical Phase II metabolism conjugation reactions

Indeed, metabolism is an extremely complicated aspect of medicinal chemistry: it challenges the development team to maintain a strong target binding affinity whilst simultaneously balancing a favourable half-life profile, minimising toxicity issues yet still accounting for the polymorphisms of individuals. Data points are difficult to obtain, as typically animal studies or clinical trials have to be used for this optimisation. Because of this, it is one of the major contributory factors of the massive costs involved in drug research.

Excretion

In humans, there are three major routes available for a drug to leave the body: the urine, faeces, and breath. The most common route, *via* urine, occurs when chemicals have sufficient hydrophilicity to be removed from circulation by the kidneys. Some drugs are actively transported into the bile, which in turn is secreted into the gastrointestinal tract. Also, compounds which have very low molecular weight may be gaseous and these are exchanged at the lungs to leave the body.

Per regulatory requirements, it is the responsibility of the drug company to show that the drug and all its metabolites are completely excreted from the body. Failure of this could once again lead to toxicity issues. This can be studied by the use of radio-labelled drugs which can be traced as they leave the body.

Toxicity

Any compound in excess is toxic and as such, toxicity issues are essentially unavoidable. This is due to the unplanned, promiscuous activity of drugs (or their respective metabolites) on unintended substrates. As long as the drug can exert its beneficial effect before it becomes toxic, it is deemed safe. To quantify this safety margin, both the ED₅₀ and TD₅₀ dose response curves are attained (Figure 1-7), where ED₅₀ is the dose at which 50% of patients experience the beneficial aspect and TD₅₀ is the dose at which 50% of patients exhibit toxic side effects.

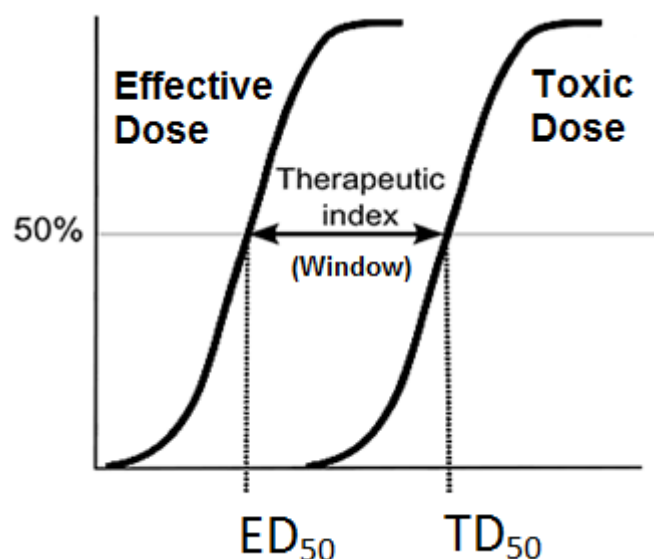


Figure 1-7 Graphical representation of the therapeutic window.

Following this, the therapeutic index (or therapeutic window) can be calculated:

$$\text{Therapeutic index} = \frac{\text{TD}_{50}}{\text{ED}_{50}} \quad \text{Eq 5}$$

A lower ED₅₀ and a higher TD₅₀, leads to an improved (larger) therapeutic index. Anything above 50 is considered to be a relatively safe drug.¹⁸ If the therapeutic index becomes too low, the drug could be withdrawn. However, depending on the availability of competing products and the seriousness of the ailment to be treated, active monitoring of drug levels in the body can allow certain drugs to remain on the market even though they have a very small therapeutic window.

There are several ways to increase this window: Firstly, the drug's ED₅₀ can be improved by increasing the affinity for the target. The drug may also be administered in an inactive form as prodrug. The prodrug may then have reduced toxicity issues, and is then converted in the body to its active form. For example, aspirin is sometimes administered as a prodrug. The carboxylic acid is known to cause gastric problems. By masking it as an ester, this unwanted issue is avoided (see Scheme 1-1).^{54, 55} Prodrugs are relatively common: 16% of drugs on the market are prodrugs.²⁸

Scheme 1-1 Ester prodrug of aspirin hydrolysing to the bioactive pharmaceutical – acetylsalicylic acid



Another way is to localise the active drug to the site of action, thereby reducing the systemic concentration of the drug. Targeting can be exploited by using the selectivity offered by antibodies. Antibodies have strong affinities for larger chemical structures which may otherwise be difficult to target. By covalently linking the drug to an antibody, the drug becomes localised in the vicinity of the antigen (target of the antibody).⁵⁶ This strategy is particularly prominent in cancer chemotherapy.⁵⁷ If it is known that the drug's metabolites are the reason of concern, the drug can be engineered to yield only non-toxic metabolites. The drug is then termed a "soft drug".^{58, 59}

In the initial phases of development, there are two critical tests which drug candidates simply must not fail. The first concerns mutagenicity. To check for this property, the Ames test has been developed.⁶⁰ Also, compounds which may interfere with hERG potassium ion channels, involved in the regulation of heartbeat, are also quickly redesigned or discarded.⁶¹

Bioisosterism

Bioisosteres are replacement moieties which are used in the fine tuning of molecules to produce more favourable pharmacodynamic or pharmacokinetic results. They are chosen to occupy similar spatial and electronic arrangements of important functional groups, yet impart subtle differences on the molecule. Amongst other things, these alterations can lead to improved permeability, increased or decreased metabolic stability, conformational entropy gains, lowered toxicity. Examples include substituting hydrogen for fluorine to decrease metabolic susceptibility; carbon for silicon to allow geminal diols; carboxylic acid for tetrazole

or squaric acid to avoid formal charges. A comprehensive review on this topic has been published.⁶²

Clinical trials and moving from bench to bedside

Testing compounds *in vitro* can only be done up to a point and human testing is then required. Clinical trials do this in a systematic fashion, always keeping patient safety at the forefront. Drugs which enter the trials undergo rigorous testing and clearing each of the 3 phases in the trials signifies a major milestone in drug development. It is important to note that only 11% of drug candidates which enter the trials succeed, so the financial risks are significant in this extremely expensive process (see Figure 1-8).¹⁰ However, for each phase cleared, the value of the drug also increases vastly. This allows earlier investors exit strategies (a point at which invested resources can be recuperated).

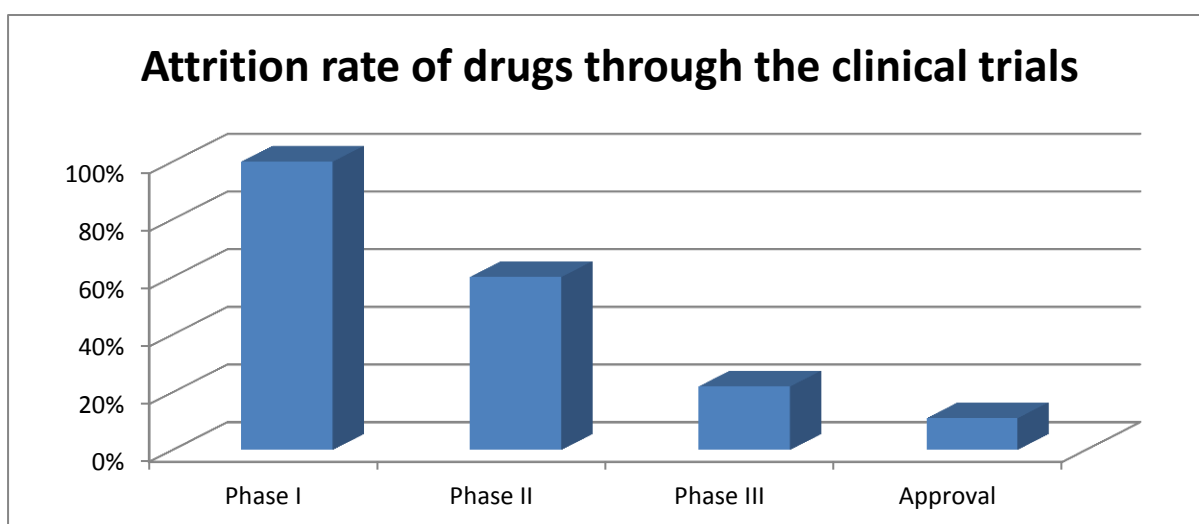


Figure 1-8 Histogram indicating how common it is for drugs to fail in the clinical trials.¹⁰

In the special pre-phase I of the trials, patients are administered mere pin pricks of the samples to check for any severe interactions the drug may have. Following this are the phase I tests. During this phase, healthy patients are administered the drug. Properties which are tested in this phase include: establishment of toxicity profiles, drug half-lives, how the drug is metabolised and excreted. Should this succeed, phase II commences.

In phase II, the drug candidate is administered to sick patients and monitored to see how well they respond to treatment. Following this, dosing requirements are established. The number of patients is also increased.

In phase III, the same protocols are followed as in the phase II although significantly more patients are tested for rare side effects. This also increases statistical confidence in the findings.

Chapter 1 – Introduction

Should phase III succeed, huge celebrations are in order: The drug is registered at relevant regulatory bodies and is placed on the market. During this time, the so called phase IV trials maintain a watch on how the drug performs. This includes how patients react who use the drug long term and for very rare side effects. Should serious consequences arise, the drug may be withdrawn from the market with little notice.

As a result of the large costs involved in the development of diseases, pharmaceutical companies have learned that it is financially less viable to treat diseases of the Third World as patients who can ill-afford food are less likely to afford medication. This has given rise to the presence of the so-called “neglected diseases”. These include AIDS, tuberculosis and malaria. For the most part, non-profit organisations and academia attempt to combat these diseases.

2. Malaria

Introduction and life cycle

The latest World Health Organisation Malaria Report estimates the number of malaria deaths to be below 1 million, with the figure quoted as 660 000 in 2010. 90% of the cases occur within poverty stricken of sub-Saharan Africa, a region which also accounts for 75% of the fatalities. Even more disconcerting, malaria is the cause of 20 – 25% of child deaths in the region.^{63, 64} There is also a complex relationship between the spread of malaria and the socioeconomic burden associated with it, as the one sustains the other.^{65, 66}

The malaria disease is an infection of the protozoan parasite of the *Plasmodium* genus, vectored by the female *Anopheles* mosquito. In the *Plasmodium* genus, four subgenera are mainly responsible for 90% of malaria infections: *P. vivax*, *P. ovale*, *P. malariae* and *P. falciparum*. The latter is the agent of the most rampant and treatment resistant form of malaria.⁶³ After a carrier mosquito infects a person, the sporozoites localise to the liver, where they infect hepatocytes. In *P. vivax* and *P. ovale*, some of the sporozoites become dormant hypnozoites, raising the chance of a relapsing infection. However, for all four genera, these sporozoites become actively growing schizonts. When these schizonts are grown, they lyse the host cell and release merozoites. These infect host erythrocytes, dividing into macro and micro gametocytes, to begin the sexual stage of the lifecycle. It is during this stage that, once ingested by another mosquito, the male and female gametocytes fuse to form diploid ookinetes in the gut of the mosquito. These ookinetes are then able to form oocysts by entrenching themselves in the gut cells. From here, meiotic division forms fresh sporozoites to begin the cycle anew (see Figure 2-1).⁶⁷

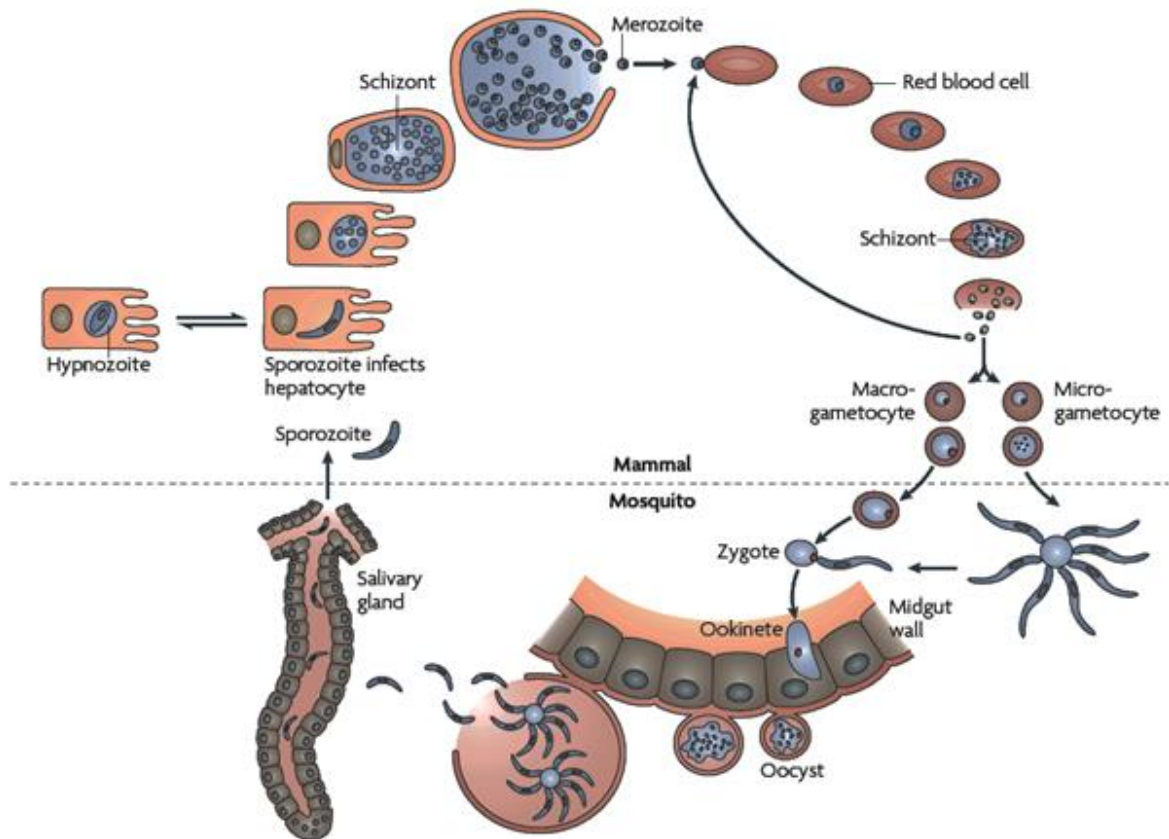


Figure 2-1 Lifecycle of Plasmodia species.⁶⁷

Current antimalarials and their mechanisms of action

As of 2012, no vaccine has successfully been developed against the parasite. Therefore, the primary method of fighting the disease entails the use of small molecules, of which there are seven major classes.

4-Aminoquinolines

As the protozoa metabolise hemoglobin from host cells, a toxic by-product, known as ferriprotoporphyrin IX (FPPIX), is released within the parasitic food vacuole. To remove the toxic FPPIX from solution, the parasite crystallises FPPIX to non-toxic crystals, known as hemozoin, by a pi-stacking mechanism. The 4-aminoquinolines inhibit this detoxification process of crystallisation by an unknown mechanism. However, it is believed that the 4-aminoquinolines, which are concentrated within the food vacuole by an unknown mechanism, inhibit the crystallisation process by complexing to hemozoin.^{68, 69} This class of compounds, championed by chloroquine (CQ), were introduced at the end of World War 2. Initially, they were extremely effective and were believed to send malaria “well on its way towards oblivion”.⁷⁰ Unfortunately, as with most other antimalarials, resistance has become very common in the early 21st century, with resistance reported in 80 – 100% of cases

worldwide.^{71, 72} Several genes have been identified to be involved in 4-aminoquinoline resistance. The gene *pfmdr* encodes the protein for the drug efflux pump.^{73, 74} The protein product of *pfcr1* is involved in the removal of CQ from the food vacuole.^{73, 75} Although *pfcr2* had been considered to be involved in resistance, this notion has been discredited.⁷⁶ Drugs falling under this class include: Chloroquine and Amodiaquine (see Figure 2-2).

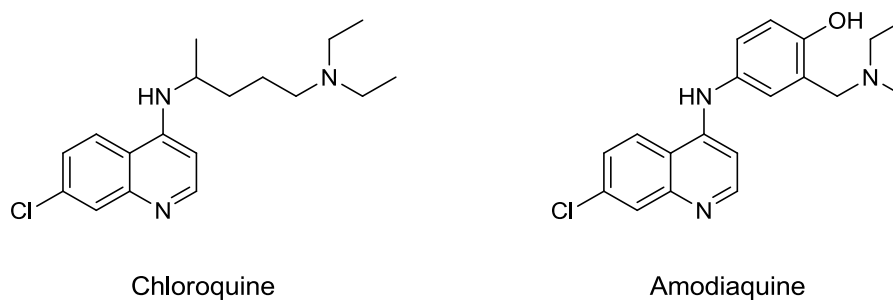


Figure 2-2 4-Aminoquinolines

8-Aminoquinolines

This class of antimalarials are only used for radical cure (removal of hypnozoites) of *P. vivax* infections. Suspected to interfere with the Golgi apparatus' activity by inhibition of vesicle formation.⁷⁷ Resistance mechanisms are currently not characterised. The only 8-aminoquinoline used clinically is Primaquine (see Figure 2-3).

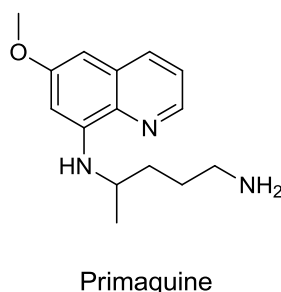
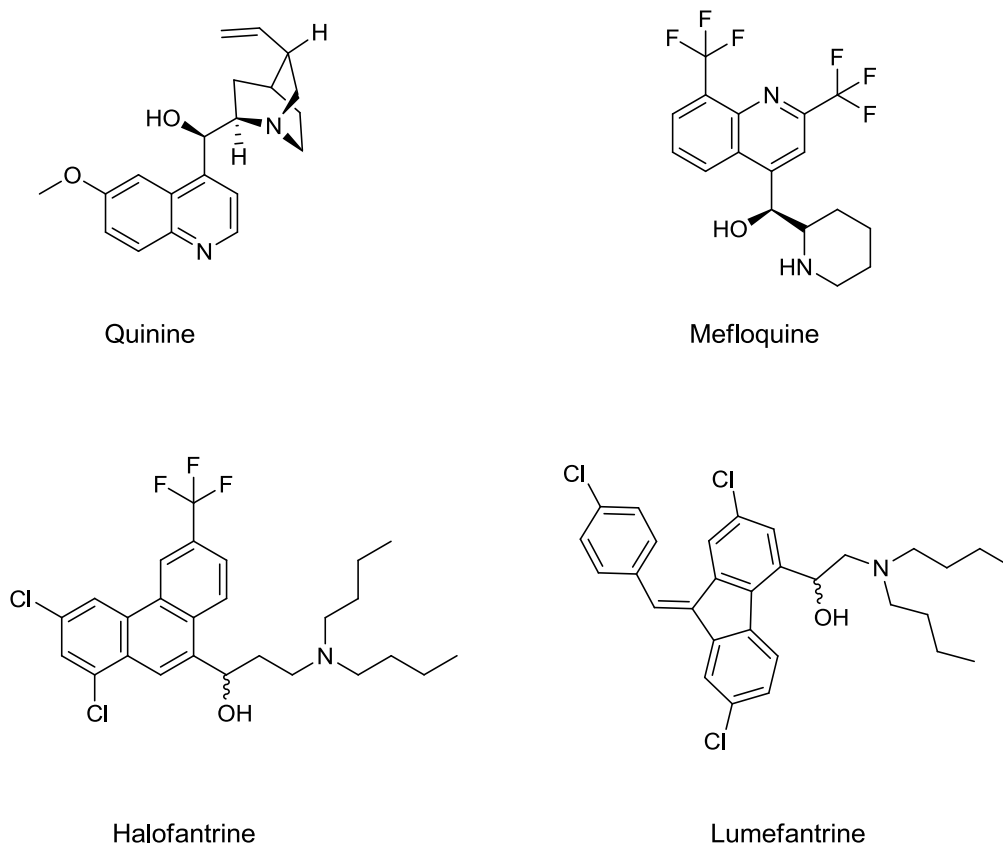


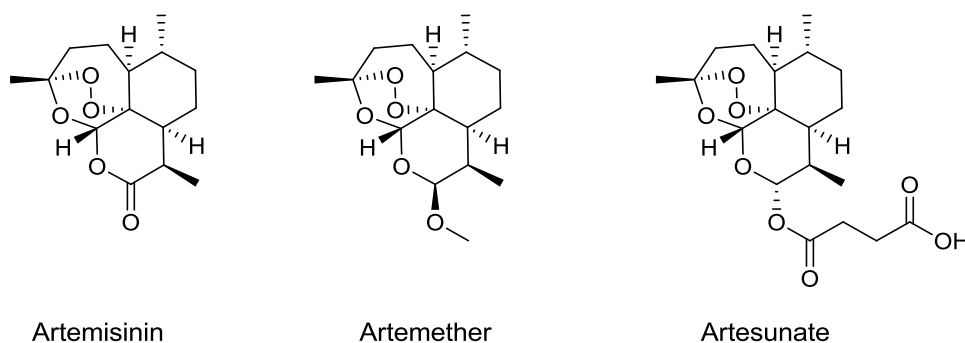
Figure 2-3 8-Aminoquinolines

Arylaminoalcohols

This class of compounds is suspected to interfere with the heme ingestion mechanism^{78, 79} The mechanism of resistance is once again the protein efflux pump protein, pfMDR1 protein, which is encoded by the *pfmdr1* gene. Resistance is found in Southeast Asia and Western Oceania.⁸⁰⁻⁸³ Drugs falling under this class include: Quinine, Mefloquine, Halofantrine, Lumefantrine (see Figure 2-4).

**Figure 2-4 Arylaminoalcohols****Artemisinin**

The artemisinin derivatives are the current gold standard in antimalarials, characterised by rapid parasitic clearance times from the body. Unfortunately they have very short half lives in the body which limits their effective duration. The artemisinin's (see Figure 2-5) exact mechanism of action is currently uncertain, although it likely entails the disruption redox homestasis within the cell. Furthermore, the endo-peroxide is essential for activity and that it is affected by a radical-type mechanism.⁸⁴⁻⁸⁷ Similarly to other classes of antimalarials, this class is exclusively used in combination therapy to limit the potential for resistant strains to develop, typically with drugs of longer duration.⁸⁸ Should the parasite develop resistance to artemisinin, the second antimalarial would kill the parasite. In this case, artemisinin is "protected". Unfortunately, the reverse does not hold true if the artemisinin has already been cleared from the body. There have been several alarming reports of reduced sensitivity in some parts of the world.⁸⁹⁻⁹¹ Resistance is once again bestowed by the *pfmdr1* gene,⁸⁰⁻⁸² as well as a mutation in the *pfatp6* gene.^{92, 93} Drugs falling under this class include: Artemether, Artesunate.



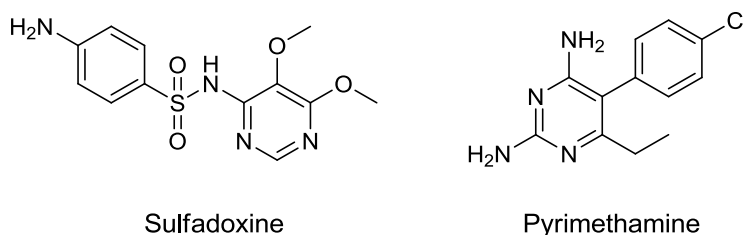
Artemisinin

Artemether

Artesunate

Figure 2-5 Artemisinin and derivatives**Antifolates**

The antifolates act as competitive inhibitors towards two enzymes which are both involved in the folate biosynthetic pathway: Sulfadoxine and Pyrimethamine (see Figure 2-6) are administered as combination therapy due to their synergistic effects in that they target enzymes in the same pathway: dihydropteroate synthase and dihydrofolate reductase, respectively.⁹⁴ The advantage of having multiple targets is that both targets need to mutate at the same time for resistance to develop, which is significantly less likely. As with all other anti-malarials, resistant strains have been reported. Resistance is conferred by point mutations in the genes encoding for the two enzymes, *pfdhps*⁹⁵ and *pfdhfr*.⁹⁶

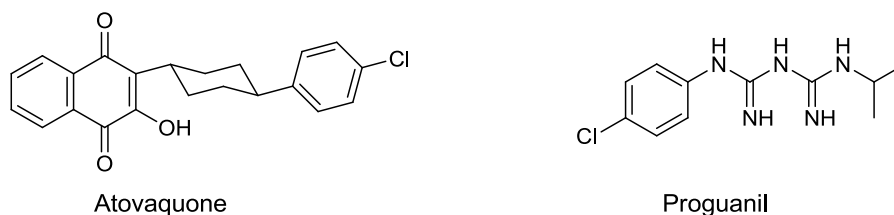


Sulfadoxine

Pyrimethamine

Figure 2-6 Antifolates**Respiratory chain inhibitors**

By binding to the cytochrome bc1 complex⁹⁷ the drug inhibits the mitochondrial electron transport chain, leading to a loss in the membrane potential.⁹⁸ The Y268S mutation reduces affinity for the drug for the binding site.^{99, 100} Drugs falling under this class include: Atovaquone, Proguanil (see Figure 2-7).



Atovaquone

Proguanil

Figure 2-7 Respiratory chain inhibitors

Antibiotics

The tetracyclines are primarily antibiotics, but also exhibit anti-malarial activity. They interfere with translation by binding to the ribosomal subunits which are involved in translation.¹⁰¹ Tetracycline resistance in malaria is currently uncharacterised. Other drugs falling under this class include: Doxycyclin, Clindamycin (see Figure 2-8).

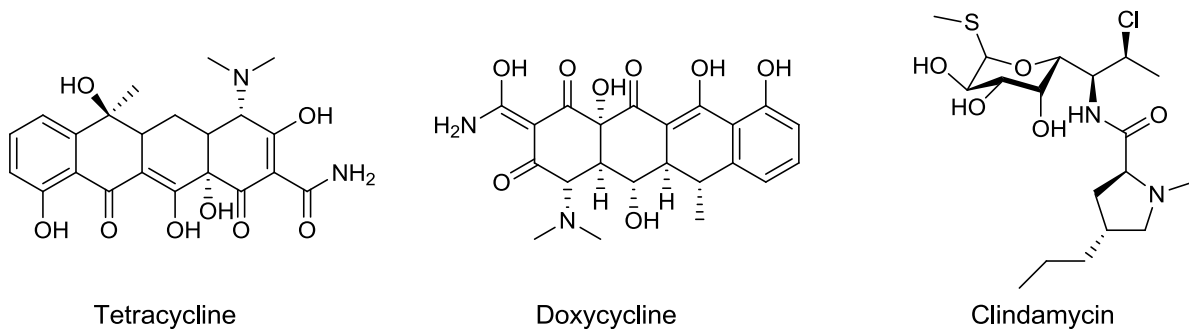


Figure 2-8 Antibiotics

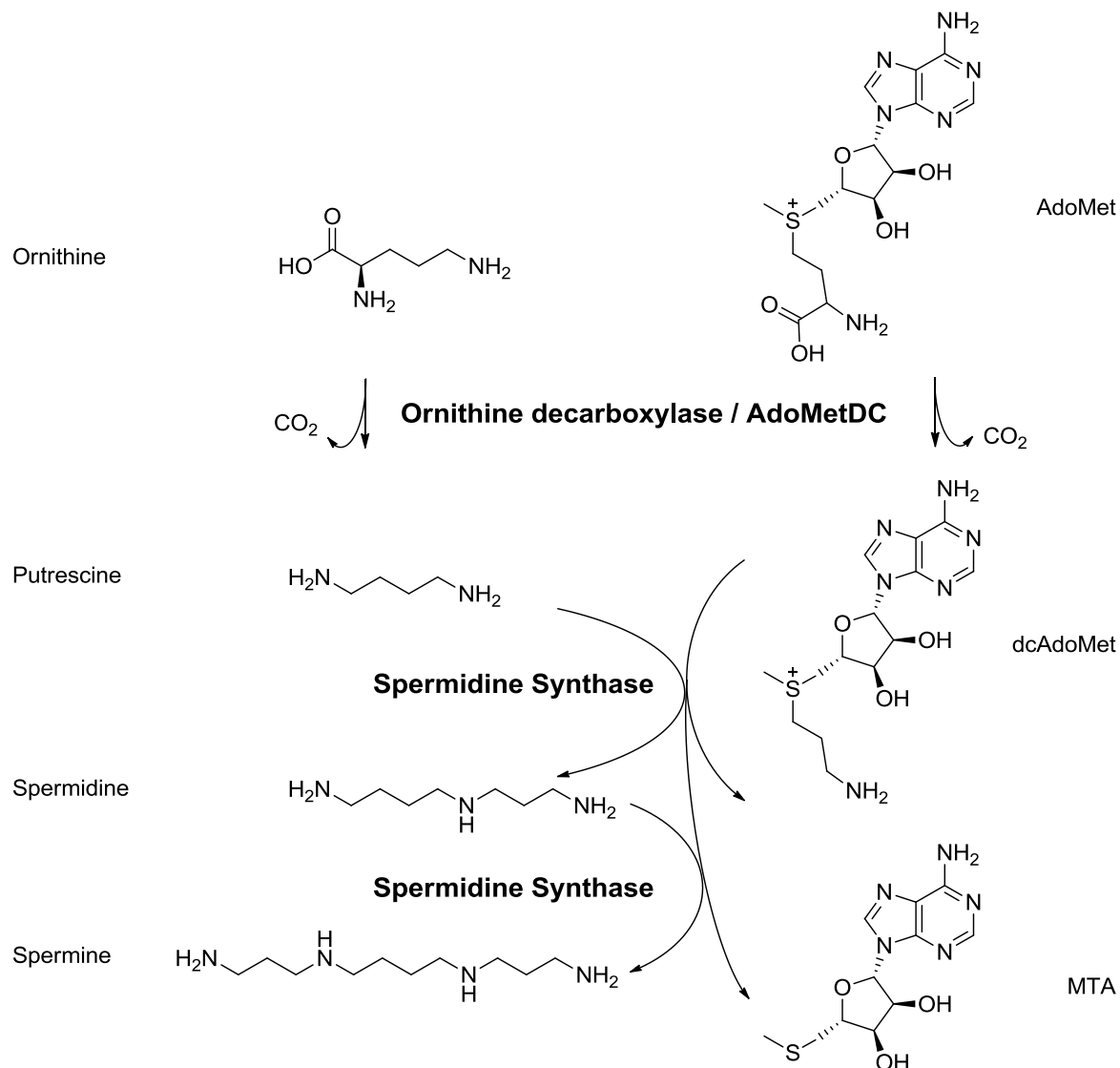
The biology of *Plasmodium* parasites has been well studied and numerous pathways have been identified to be essential for the parasite's growth – all of which contain putative drug targets. Among many others, these include: glycolysis, biological methylation, Coenzyme-A biosynthesis, and purine metabolism.¹⁰² In our efforts to combat the disease, we have identified two different enzymatic targets for inhibition: spermidine synthase from the polyamine biosynthetic pathway and protein farnesyltransferase, which is involved in post-translational protein modification.

Inhibition targets for this dissertation

Polyamine biosynthesis - Spermidine synthase

The polyamines (putrescine, spermidine, spermine and cadaverine) seen in Scheme 2-1, are small molecules which are involved in a number of important cellular processes. Being highly flexible and cationic at pH 7.4, they have numerous interactions with anionic moieties within the cell.¹⁰³ These interactions include DNA, RNA, ATP, certain membrane bound proteins¹⁰⁴⁻¹⁰⁷ and phospholipids.¹⁰⁸ Furthermore, the polyamines are also known to be involved in transcriptional regulation via eukaryotic initiation factor 5 (eIF5).¹⁰⁹ Due to this ubiquity, the polyamine biosynthetic pathway is an essential part of cellular metabolism and is therefore a viable target for inhibition.¹⁰⁹

In mammals, the first half of the polyamine biosynthetic pathway (see Scheme 2-1) initiates with the decarboxylation of L-ornithine by ornithine decarboxylase (ODC) to produce 1,4-diaminobutane (putrescine).¹¹⁰ The other half of the pathway is catalysed by S-adenosylmethionine decarboxylase (AdoMetDC). This enzyme decarboxylates S-adenosylmethionine (AdoMet) to produce decarboxylated S-adenosylmethionine (dcAdoMet).¹¹¹ However, in *P. falciparum* these two enzymes are fused as a unique, hinge joined bifunctional protein, ODC/AdoMetDC.¹¹² Next, the pathway converges by the action of spermidine synthase (SpdSyn). SpdSyn transfers the aminopropyl moiety from dcAdoMet to putrescine to produce spermidine and the by-product, 5'-methylthioadenosine (MTA). Additionally, spermine synthase (SpmSyn) is able to transfer a second aminopropyl from another dcAdoMet to produce spermine and MTA.¹¹³ However, in *P. falciparum* this enzyme is not present and spermine production is also catalysed by SpdSyn.¹¹⁴ Inhibitors for all of these enzymes have been developed.¹¹⁵⁻¹¹⁷ For the purposes of this dissertation, we have focussed our efforts on the inhibition of pfSpdSyn.

Scheme 2-1 Polyamine biosynthetic pathway in *P. falciparum*

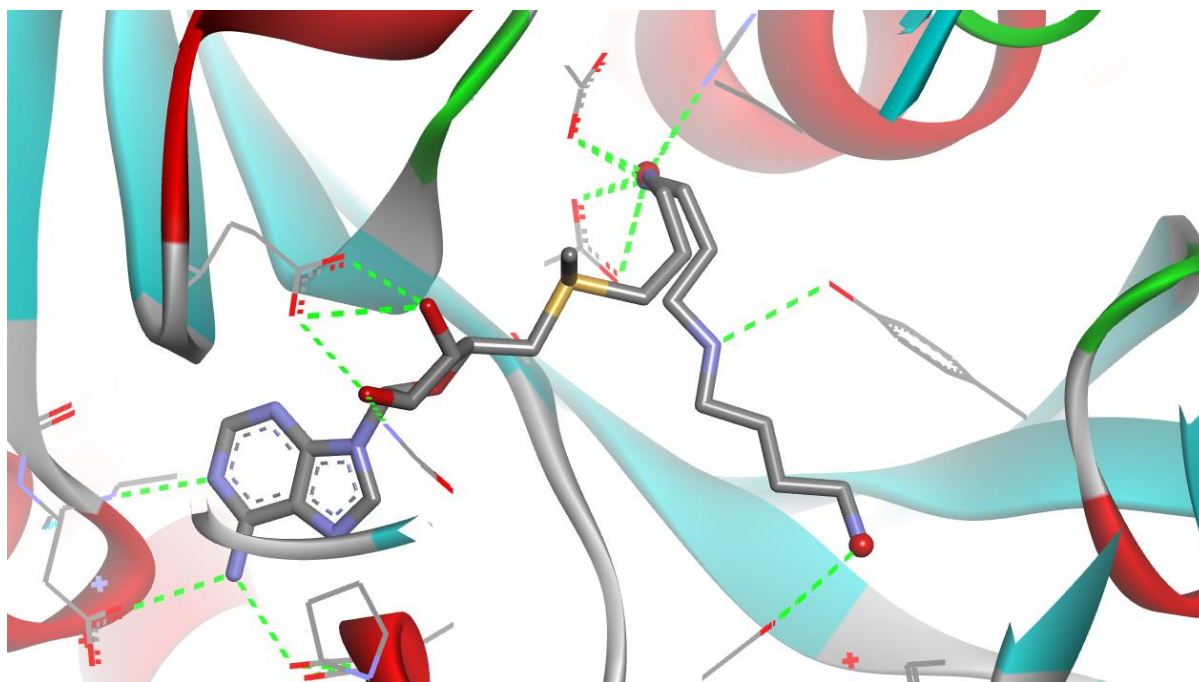
The validity of SpdSyn inhibition has been demonstrated in several species, through the use of SpdSyn null mutants: *Dictyostelium discoideum*,¹¹⁸ *Leishmania donovani*^{119, 120} and *Aspergillus nidulans*.¹²¹ Although a similar null mutant has not been obtained for *P. falciparum*, the use of SpdSyn inhibitors has been shown to inhibit the parasite's growth. Moreover, the use of the same inhibitors did not show notable effect on the growth in mammalian cells.¹¹⁴

The enzyme SpdSyn consists of two homodimers, each sized 36.6 kDa with a surface contact area of 1460 Å². The K_m values for its two substrates, dcAdoMet and putrescine, are 35 and 25 μM respectively.¹¹⁴ A number of crystal structures of this enzyme, co-crystallised with inhibitors, have been published and the notable excerpts are tabulated in Table 2-1.

Table 2-1 Important crystal structures of SpdSyn

Ligand	PDB ID	Resolution (Å ²)	K _i	IC ₅₀	Reference
<u>Natural ligands</u>					
Apo	2PSS	2.2	-	-	122
MTA	2HTE	2.0	-	-	123
Spermidine	2PWP	2.1	-	-	Unpublished
Spermine	3B7P	2.0	-	-	Unpublished
<u>Inhibitors</u>					
4MCHA	2PT9 ¹²²	2.2	0.18 μM	35 μM	114, 124, 125
AdoDATO	2I7C ¹²²	1.7		8.5 uM	126-128

The active site (see Figure 2-9) of SpdSyn has four notable interaction regions: 1) Nucleoside binding region; 2) Proximal cation binding region where the amino propyl moiety from MTA is located; 3) Hydrophobic region around the alkyl chain of putrescine; and 4) Distal cation binding region where the far amine of putrescine is held in place.

**Figure 2-9 Active site of SpdSyn.**

Shown here is the superimposition of dcAdoMet (left) and spermidine (right). The red dot to the upper right indicates the position of the proximal cation pocket and the red dot to the lower right indicates the position of the distal cation pocket. Hydrogen bond interactions are displayed in green.

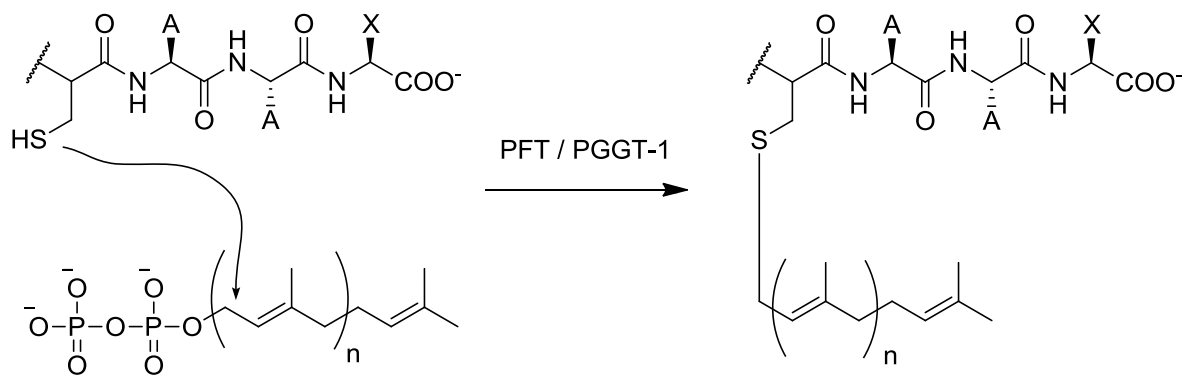
For our efforts to combat malaria, we plan to target the polyamine pathway by designing and synthesising small molecules which would have similar interaction points as the molecules which are depicted here. Additionally, our anti-malarial plan will also target a second enzyme which is involved in another essential pathway in malarial metabolism.

Protein prenylation - Farnesyltransferase

The second target for inhibition involves the enzyme which is responsible for *prenylation* – protein farnesyltransferase. Prenylation entails the post-translational conjugation of an isoprenoid lipid (prenyl) to a protein. These prenyl groups are postulated to serve two major functions: Anchorage of a protein to the cellular membrane, (proteins for these function include the Ras family and the GPCR's¹²⁹⁻¹³³) and the mediation of protein-protein interactions.^{134, 135} It is estimated that 0.5% of proteins are prenylated, demonstrating that disruption of this ubiquitous protein modification could have significant repercussions throughout the cell.¹³⁶

There are two types of isoprenoid lipids: farnesyl and geranylgeranyl, the latter having an additional 5 carbon monomer (see Scheme 2-2). The isoprenoid lipids are sourced from the mevalonic acid pathway and are found as their corresponding diphosphates (farnesylpyrophosphate and geranylgeranylpyrophosphate).¹³⁷ These isoprenoids are conjugated to proteins through the action of the enzymes Protein Farnesyltransferase (PFT) or by the two geranylgeranyltransferases (GGT1 & GGT2).¹³⁸⁻¹⁴² In contrast to mammalian cells which also have these latter two enzymes, *P. falciparum* relies solely on PFT to exert this critical prenylation.¹³⁹ For this reason, the parasite is significantly more susceptible to prenylation inhibition when compared to mammalian cells - this forms the basis of selectivity.¹³⁹ For this reason, PFT can be defined as a target for inhibition and is the focus of the second part of this dissertation.

PFT recognises (with some promiscuity) the amino acid sequence CA₁A₂X as substrate (where C = Cysteine, A₁ & A₂ are aliphatic residues and X is any C-terminal amino acid). Farnesylation occurs at the thiol group from the cysteine residue and in the process the pyrophosphate acts as a leaving group.¹⁴³

Scheme 2-2 Reaction mechanism of protein farnesylation

n =2 Farnesyl pyrophosphate

n =3 Geranylgeranyl pyrophosphate

PFT is a heterodimeric metalloenzyme, consisting of an α and β subunit, sized 48 and 46 kDa, respectively. It has a very large active site to accommodate both the amino acid sequence and the farnesyl moiety. At the time of writing, there were 55 PFT structure deposits in the PDB. 41 of these are from *Rattus norvegicus* and 14 from *Homo sapiens*. Presented in Table 2-2 are the notable crystal structures of relevance to this dissertation.

Table 2-2 Notable crystal structures of mammalian PFT

Ligand	PDB ID	Resolution (\AA^2)	IC ₅₀	Reference
Apo	1FT1	2.3	-	144
PB-93	2R2L	2.2	1 nM	145

Aims

For both these enzymatic targets, we shall design novel potential antimalarial agents. These designs will be inspired by currently known inhibitors which will be optimised further by the use of published crystal structure data and modelling techniques.

Once these designs have been made, they are to be synthesised and tested for activity against the causative agent of malaria, *P. falciparum*.

3. Molecular modelling – General

Over the years, the advancement of computer technology has changed our lives in significant ways. From predicting the weather, calculating satellite orbits to our simple day to day email requirements, it has revolutionised society. Recently, computational power has reached the point where we can feasibly simulate the interface between chemistry and biology to have a direct impact in the drug development pipeline.

Molecular mechanics

To model a chemical system without compromise, the Schrödinger equation offers unparalleled accuracy using quantum mechanics (QM). Sadly, this high degree of accuracy requires an extreme degree of computational power which is simply not available for molecules much larger than ethanol. Biological systems typically contain thousands of atoms and applying first principles to these cases is completely unfeasible in the early 21st century. Efforts to reduce this high computational cost, whilst maintaining acceptable accuracy, have given rise to molecular mechanics (MM). Comparatively, MM has the advantage of significantly reduced computational requirements to QM, albeit at the sacrifice of accuracy. This is implemented through the use forcefields which make approximations on the way molecules twist, move, and otherwise interact with one another.

Forcefields and CHARMM

A range of forcefields are available, each with their own special niches which may overlap with another. For example the UFF forcefield (Universal forcefield)¹⁴⁶ is able to accommodate almost all element types, albeit at reduced accuracy as compared to a more specialised forcefield such as MMFF (Merck molecular forcefield).^{147, 148} Although MMFF does not support nearly as many elements, it offers superior accuracy when simulating organic molecules. This includes accurate portrayal of conformational energetics and non-bonded interactions. Another forcefield, CFF (consistent force field), offers better incorporation of explicit solvent models and specialises in the modelling of molecules that are in the gas phase.^{149, 150}

The CHARMM forcefield, (Chemistry at HARvard Molecular Mechanics) is the preferred forcefield in DS (Accelrys Discovery Studio). It is a general purpose forcefield that is aimed to be suitable for most systems relevant in the drug discovery process.^{151, 152}

Generally, CHARMM works as such: Firstly, the molecule is “typed”. To do this, each individual atom is examined and characterised by several properties, which include the element type, oxidation state, hybridization state, neighbouring atom type and its general

“chemical environment” (see Figure 3-1). Each atom is then assigned parameters based on these properties, to be used in the energy equation. Amongst other parameters, these include equilibrium bond length, equilibrium torsion angles, van der Waals contact radii and partial charges. Like other forcefields, CHARMM’s parameters are obtained from large tables, themselves derived from training sets of X-ray diffraction data and/or *ab initio* calculations.¹⁵¹

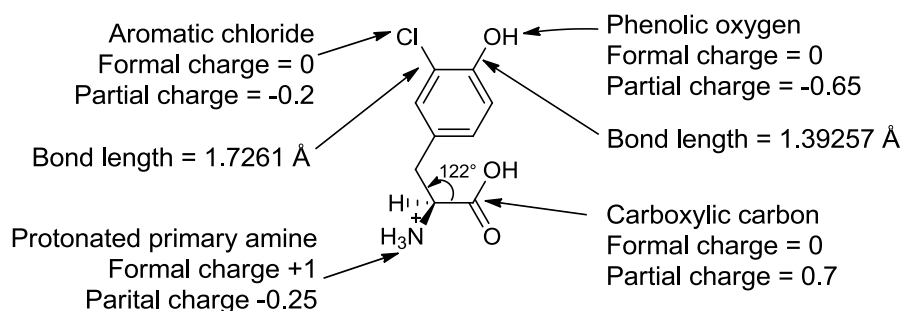


Figure 3-1 Simple depiction of how various atoms are assigned parameters¹⁵¹

To reduce the computational cost in CHARMM, atoms and bonds are modelled analogous to simple balls on springs which are influenced by Coulombic forces. In CHARMM, the total energy of a system is calculated by the sum of the individual energy functions and is given by:¹⁵¹

$$E_{\text{Pot}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{out of plane}} + E_{\text{elec}} + E_{\text{vdW}} + E_{\text{constraint}} \quad \text{Eq 6}$$

Expanded, these energetic terms are described by the following approximative equations:¹⁵¹

$$E_{\text{Pot}} = \sum k_b(r - r_0)^2 + \sum K_\theta(\theta - \theta_0)^2 + \sum |k_\phi| - k_\phi \cos(n\theta) + \sum k_\chi(\chi - \chi_0)^2 + \sum \frac{q_i q_j}{\epsilon 4\pi\epsilon_0 r_{ij}} + \sum \left[\epsilon_{ij}^{\text{min}} \left(\frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] + E_{\text{constraint}} \quad \text{Eq 7}$$

Bonded interactions

The first four terms describe the interactions between two covalently bonded atoms. These terms describe the bond length and angles between said atoms. These terms are deviations from Hooke’s law which describes two objects connected by a spring: the further away from equilibrium, the higher the energy state. The equilibrium distances and force constants (ie. r_0 & k_b for the first term, respectively) are obtained from the CHARMM parameter table files.¹⁵¹

Non-bonded interactions

The next two terms describe the non-bonded interactions between atoms in close proximity. These include the electrostatic and van der Waals interactions. Noteworthy, for reversibly binding inhibitors, these are the only interactions between a drug candidate and its target.

These non-bonded interactions make up the bulk of the computational load when modelling biological systems using MM. In an attempt to decrease the computational cost of these important interactions, parameters are included that ignore interactions above a defined cut-off distance, usually about 14 Å. However, introducing a discontinuous function (a value suddenly changing to 0) as we move beyond the cut-off distance may lead to discrepancies which could become prominent during minimisation (see below). To minimise this, an intermediate region is defined where a smoothing function is applied, approaching 0 at the outer boundary.¹⁵¹

Electrostatic term

The electrostatic interaction term, which is derived from the law of Coulomb, quantifies the energy between two point charges. It is given by:

$$E_{\text{elec}} = \sum \frac{q_i q_j}{\epsilon 4\pi\epsilon_0 r_{ij}} \quad \text{Eq 8}$$

Where q_i and q_j are the charges (defined as the partial charges by the force field), ϵ is the dielectric constant (polarizability) of the medium, ϵ_0 is the dielectric reference within vacuum and r is the distance between the charges.

Van der Waals term

Van der Waals contacts are estimated by the Lennard Jones 6-12 potential equation.¹⁵³ A limitation is that there is no directionality in this estimation, as the potential is spherically symmetric.¹⁵³

$$E_{\text{vdW}} = \sum \left[\epsilon_{ij}^{\text{min}} \left(\frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] \quad \text{Eq 9}$$

Where R_{ij} is the optimum distance and ϵ_{ij} is the well depth of the Lennard Jones potential (see Figure 3-2).

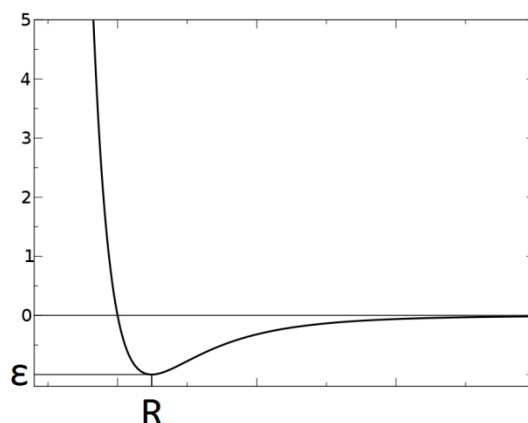


Figure 3-2 Graphical plot illustrating the Lennard Jones potential.¹⁵³

At small r , the first term dominates and a strong repulsion is experienced by the atoms. As r approaches R , the second, stabilising term dominates and the bottom of the well is reached (given by ϵ). Finally, as r is increased, the force experienced between the atoms become negligible.

Constraint

This term is only used when the user specifies custom constraints on the system.

Solvent models

To model a molecule accurately, its environment needs to be taken into account as it can have significant effects on the molecule. Firstly, the phase needs to be decided upon. The simplest way to model a molecule is in a vacuum. This is a gross over simplification for the SBDD context. The solvated state is of primary importance as the solvent significantly influences electrostatic interactions by shielding effects. Significant computational time is dedicated to compensate the shielding phenomenon accurately and the choice of the solvation model should be carefully weighed. For ease of reference, the electrostatic term is again given.¹⁵¹

$$\sum \frac{q_i q_j}{4\epsilon\pi\epsilon_0 r_{ij}} \quad \text{Eq 10}$$

No solvation model

The dielectric constant is set to 1 (that of a vacuum).¹⁵⁴ Although by far the fastest to process, the use of this option will greatly overestimate the electrostatic energies and should generally only be used when an explicit solvent model is employed (please see below).

Implicit solvent models

Implicit solvent models allow for an efficient compromise between computational expense and an accurate depiction of the molecule in a solvated state. These approximations are particularly useful when large sets of chemical space have to be explored (large proteins,

conformational searching of molecules with many rotatable bonds, etc.) in order to estimate electrostatic energies. Of the various types that exist, three are briefly discussed here.

Distance dependent dielectrics

This is the crudest implicit solvent model in DS. It is essentially as fast as using no solvent model, yet significantly more accurate. Its main applications are for occasions when very rough estimations suffice, usually when subsequent calculations are followed up with more rigorous methods. Only a single variable is introduced by varying the dielectric constant over distance and is given by: ¹⁵⁴

$$\varepsilon(r) = \varepsilon_0 r \quad \text{Eq 11}$$

Generalised Born

The Generalised Born (GB) implicit solvent model estimates the electrostatic term by replacing the traditional sum of Coulombic charges with the Born equation. ¹⁵⁵

$$G = -166 \left(1 - \frac{1}{\varepsilon}\right) \sum_{i=1}^n \sum_{j=1}^n \frac{q_i q_j}{f_{GB}} \quad \text{Eq 12}$$

Essentially, this means that the protein is covered in spheres where $\varepsilon = 2$ and in the remainder areas $\varepsilon = 80$ (dielectric of water). This implicit solvent model is suitable for simulating biological systems with a high degree of confidence, albeit at a markedly increased computational expense.

Poisson-Boltzmann

This implicit solvent model is most accurate implicit solvent model available in DS. ¹⁵⁶ Due to the high computational cost, this method is only available for single point energy calculations (not minimisation or dynamics).

Explicit solvent model

The most accurate, yet computationally most expensive, method to model a solvent. In this way, the molecule is surrounded by a host of solvent molecules (solvated) and electrostatic energies are applied to each atom individually. A dielectric constant of 1 is used for these calculations. ¹⁵¹

On a side note, when using an explicit solvent model, Particle Mesh Ewald electrostatics may be used. This technique is then able to effectively model long range electrostatic interactions without a cut-off distance in a computationally efficient manner using Fourier space.

Minimisation

Minimisation can be divided into two main categories i.e. the search for local and global energy minima. Figure 3-3 depicts a typical energy landscape in a biomolecular system where multiple energy minima and maxima exist which varies on the conformational space. Searching for a local minimum is relatively easy and a variety of methods are available in DS to traverse directly down an energy gradient to find the corresponding local energy minimum. These methods include Adopted Basis Newton-Raphson,¹⁵⁷ Powell,¹⁵⁸ Steepest Descent,¹⁵⁹ Conjugate gradient,¹⁶⁰ and Smart minimiser.¹⁵¹

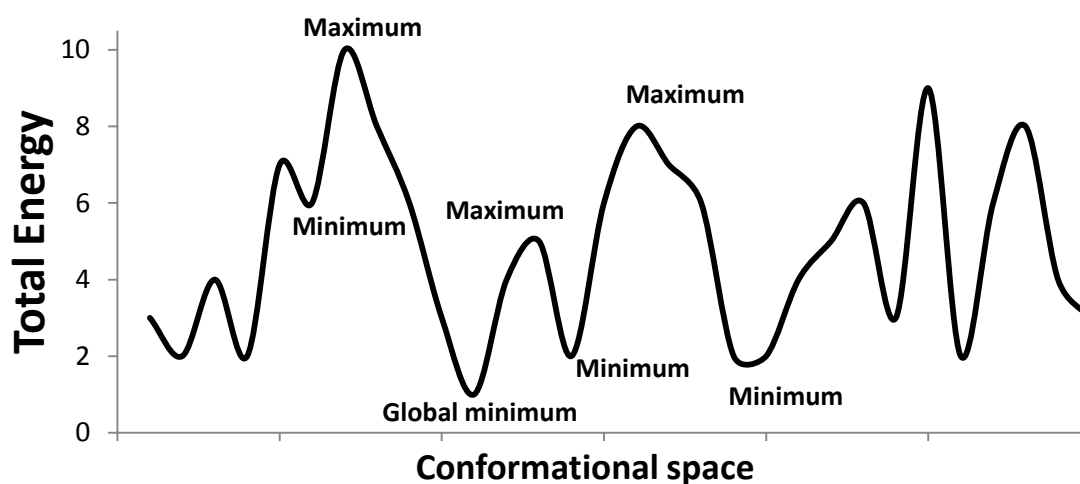


Figure 3-3 Simplified depiction of the energy landscape as function of a single variable. This is typically an oversimplification of a typical biomolecular system where there may be thousands of variables.

Searching for the global minimum of a particular system is compounded by the fact that minimisation algorithms can only proceed directly “downhill” from a given point on the energy landscape. This means that, if minimising from a random point on the energy landscape, the resultant minimum would likely only be a local minimum. In the event that the starting point was chosen directly upward from the global minimum, then the global minimum would have been found. However, given the vast number of variables in even moderately sized pharmaceuticals, this is extremely unlikely. The only way to guarantee finding the global energy minimum would be to perform a full systematic search of the entire system by doing step-wise bond rotation. This exhaustive strategy is unfeasible for anything other than the most constrained of systems due to computational expense. Although other prominent methodologies exist to aid in the search for global minima,^{161, 162} two methods are available in DS .

High temperature molecular dynamics

The temperature can be simulated by assigning random velocities to the atoms at time intervals. However, between each step of increased temperature, an equilibration step is required. This disperses the increased energy evenly throughout the molecule, according to the laws of thermodynamics. Finally, the production step allows the molecule to move for a defined time at the reached temperature. These poses can then be used by the required application.^{163, 164}

Monte Carlo simulated annealing

This method combines Monte Carlo searching with simulated annealing (otherwise known as the Metropolis method). Monte Carlo methodology involves using multiple starting points over numerous iterations to find an answer. Simulated annealing then minimises those iterations (akin to metal being cooled or *annealed*) to search for the corresponding local minima, where after energy is inserted into the system to get the system out of the energy well. The simulation is once again minimised (annealed) and the cycle restarts. After each successive iteration, the amount of energy introduced into the system is decreased. After numerous iterations, the global energy minimum is likely to be closely approached, although this can never be confirmed.¹⁶⁵

Docking

Docking entails taking proposed candidate small molecules, which could have active site complementarity, and placing them within the active site of a target protein/enzyme. Throughout this dissertation, docking was extensively used to determine the validity of all molecules that were synthesised or planned to be synthesised.

The CDOCKER (CHARMm-based DOCKER) protocol uses a combination of high temperature molecular dynamics and simulated annealing to dock molecules within a defined active site (Figure 3-4).^{166, 167} Firstly, ligand conformations are generated using high temperature molecular dynamics. Then one of two methods is used to dock the ligands: Either the full potential or a grid based method. In the grid-based method, the active site is prepared by setting up a grid (points spaced typically 0.5 Å apart). Each grid point stores the two non-bonding parameters that are notable in guest-host interactions: electrostatic potential and van der Waals clashes. This is followed by translating the generated conformers around the active site. Rigid body rotations on the bonds are performed to search for complementarity. Then, MD-based simulated annealing is performed within the active site. Finally, the placed ligand is minimised within the protein active site. Between each of these steps, internal scoring functions are applied to only progress the more

successful candidates to the next step. As the grid-based method is markedly quicker with minimal reduction in successful docking, only the grid based method was used in this dissertation.

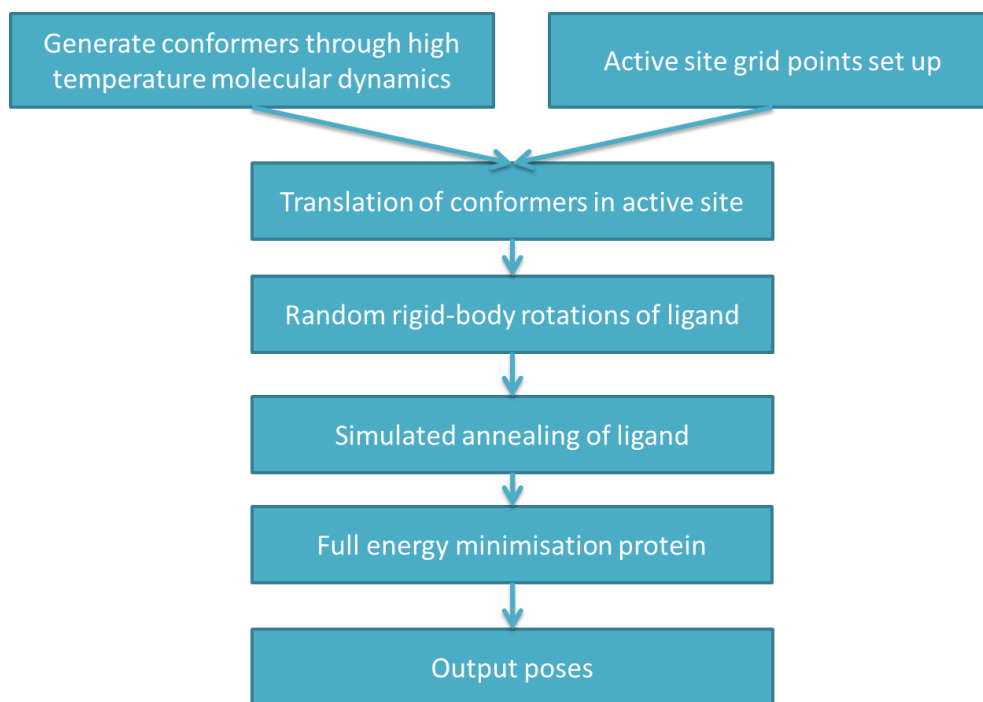


Figure 3-4 CDOCKER protocol workflow¹⁶⁶

Virtual high throughput screening and scoring

Virtual high throughput screening (vHTS) is one of the methods available in modern hit discovery efforts. It allows many millions of compounds to be screened *in silico* for complementarity to a target site at a fraction of the cost. The idea of vHTS is to start with as large a compound list as is feasible and then successively eliminate as many negatives (compounds which would have no notable interaction) as quickly as possible while gradually increasing the computational complexity with each step.

Databases

Before a vHTS can be performed, a suitable database of the molecules to be screened must be selected and prepared appropriately. This includes, amongst other things, protonation state at relevant pH, removal of molecules which are not synthetically feasible or that do not satisfy the Lipinski's rule.³⁹ Examples of some databases are listed in Table 3-1.

Table 3-1 Molecular databases

Database Name	Approximate size	Reference
Pubchem	31 million compounds	168
ChemDB	7.1 million compounds	169
Sigma-Aldrich	14 000 compounds	170
Zinc database	21 million compounds	33

An essential aspect of vHTS includes the generation of the conformations of the various molecules, known as conformers. Each single conformer is treated like an entirely different molecule during the screening process. These conformers can either be calculated on-the-fly during the screening process or be pre-generated. The latter strategy requires significant amounts of storage and CPU time to generate the conformers, but once performed this method significantly reduces screening time. This method is especially useful if the same database will be used more than once.

Docking Algorithms

Firstly, the target site of the structure is modelled and typed using a method that is specific to the vHTS algorithm. The algorithm translates and rotates each conformer through the generated pharmacophore until pre-determined satisfaction criteria are met (these criteria are dependent on the algorithm). This conformer is then saved in that position and saved as a pose for later review. Should satisfying criteria not be met, the conformer is discarded. Two of the methods available in DS are LibDock¹⁷¹ and traditional pharmacophore based screening.

Using LibDock, the active site cavity is filled with hotspots (see Figure 3-5) which are either classified as polar or non-polar. The protein is then removed and this map of hotspots is then used as a simple pharmacophore. Candidate conformers are then matched to the hotspots with satisfying criteria based on the number of hotspots in close proximity.

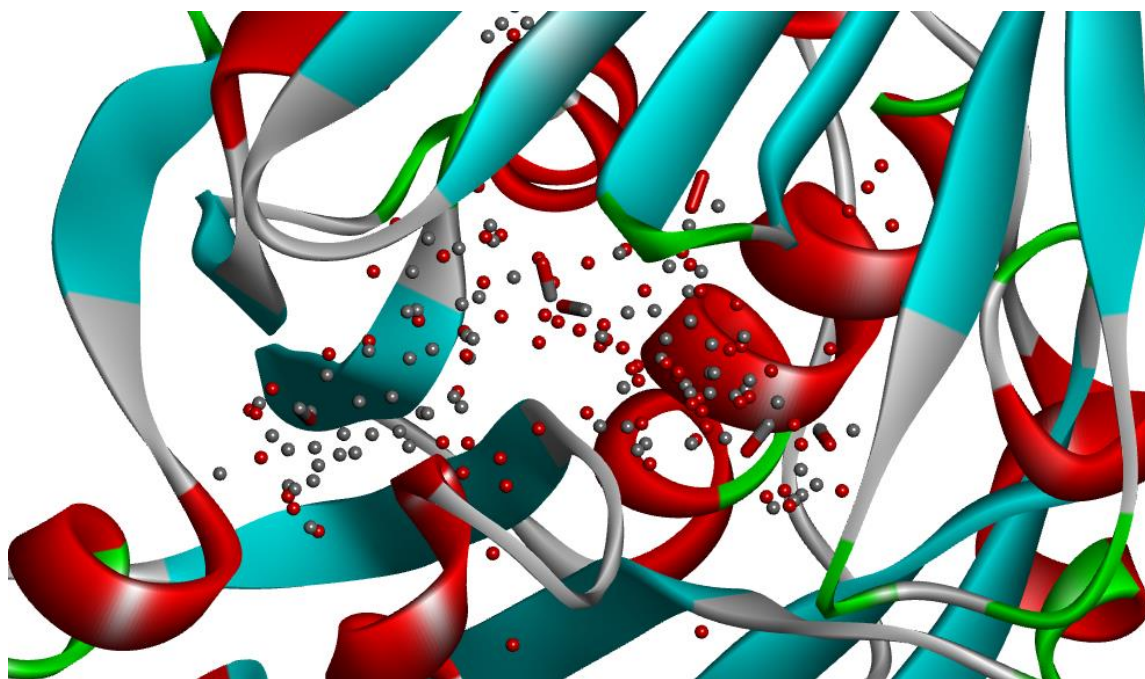


Figure 3-5 Example of hotspots as generated and used by LibDock¹⁷¹

The hotspots (seen in red and grey for polar and non-polar regions, respectively) are superimposed onto the protein structure of SpdSyn. Notice that the hotspots fall inside the cavity areas of the protein. Candidate conformers are required to align with 1 or more of these hotspots to not be discarded in a LibDock screen.

Traditional pharmacophore screening entails screening conformers through features which include, amongst other things, steric exclusions, hydrogen bond acceptor and donor features, pi-stacking interactions and hydrophobic regions.¹⁷² Satisfying criteria may include matching any combination of these things.

Scoring

Following a successful screen, many millions of potential compounds may be reduced to a few thousand. More rigorous calculations may now be performed on this smaller subset. At this stage, scoring functions are applied to vHTS results. There are many scoring functions available, each using different methods to quantify resulting poses (see Table 3-2). Because the field of scoring is still relatively new, the results from these functions may differ wildly and it is difficult to know which scoring function is the most suitable to the particular task at hand. To ascertain this, a training set of known inhibitors is used as positive control. Scoring functions that rank the positive controls high are then also used to score the potential new hits. In the unpreferred event where positive controls are unavailable, consensus scoring may be used. In this case, multiple scoring functions are applied and only poses that consistently rank top scores in these scoring functions are retained.¹⁷³

Table 3-2 Scoring functions available in DS

Scoring function	Parameters taken into account for scoring	Reference
JAIN	Lipophilic interactions Polar attractive interactions Polar repulsive interactions Solvation of the protein and ligand Entropy term for the ligand	174
LIGSCORE	Van der Waals interactions (Lennard-Jones) Buried polar surface area Total polar surface area	175
LUDI	Hydrogen bond interactions. Lipophilic interactions. Ionic interactions Entropy term for the ligand	176, 177
PLP	Hydrogen bond interactions	178, 179
PMF	Pairwise interatomic interactions Includes metal and halogen potentials	180, 181

It is important to note that in most instances, all the algorithms mentioned in this chapter will produce results. However, it is the task of the modeller to inspect these results for viability and to rerun the protocols with altered parameters or even discard the results if deemed unreliable. An ideal algorithm should be sufficiently robust to be able to accurately mimic any system. Sadly, this is not the available. The best way to check whether a simulation is valid or not, would be to test it with a positive control and applying a healthy dose of common sense.

4. Molecular Modelling - Spermidine synthase

Our antimalarial developments commenced by using molecular modelling techniques to guide us in our search for inhibitors for the enzyme spermidine synthase (SpdSyn), which was described in Chapter 2 to be a valid drug target. Among the many SpdSyn structures in the PDB, we chose to work from two structures. These two were co-crystallised with the known inhibitors 4-methyl-cyclohexylamine (4MCHA) and adenosyl-(5')-1,8-diamino-3-thiooctane (AdoDATO), respectively (see Figure 4-1).¹²²

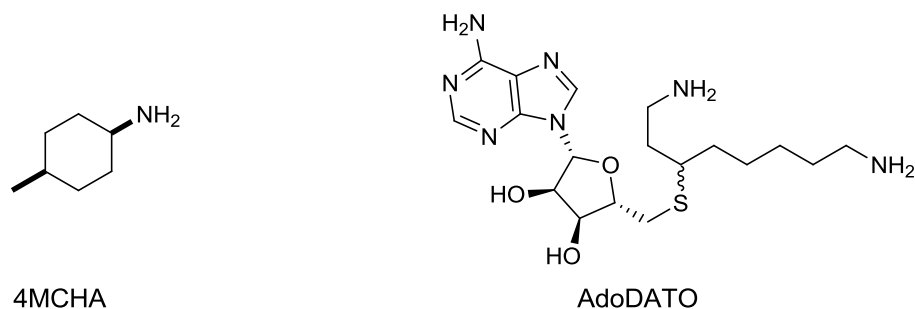


Figure 4-1 4MCHA and AdoDATO

4MCHA is much smaller and only has two notable interactions: The hydrophobic region around the cyclohexane ring and amine. AdoDATO has numerous interactions with the enzyme and also occupies a significantly larger portion of the enzyme's active site. Although it has fewer interaction points, 4MCHA is the stronger inhibitor. This is because of the significant desolvation energy penalty which AdoDATO must overcome to be inside the protein. On the other hand, 4MCHA is not penalised to the same extent. To better understand these processes, we used molecular modelling techniques.

2PT9

Our first modelling effort stemmed from the co-crystal structure of SpdSyn and 4MCHA (PDB ID: 2PT9). Unfortunately, structure files straight from the PDB are not ready for modelling purposes. As such, one of the first steps in molecular modelling is to prepare these protein models appropriately.

Protein preparation

Protein structures which are solved by X-ray diffraction (XRD), such as 2PT9, have several hurdles which need to be overcome. Firstly, protein XRD resolution, typically within the range of 1.5 – 3.0 Å, is not sufficiently accurate to identify hydrogen atoms. The positions of hydrogen atoms, including those which are involved in important hydrogen bonding interactions around the active site, have to be determined. Secondly, in the crystallisation process, the molecule is in the solid state and crystal packing may contort the protein.

Processing of the structure needs to be performed to simulate the protein in the solvated state. Next, there may also be foreign molecules which are added to aid the crystallisation process. These have to be removed manually. Finally, some protein crystal structures may have disorder in some parts of the residues. Should alternate conformations be found, the most suitable one is to be selected. Should residues be absent due to disorder, the residues (up to 20) may then be calculated back into the protein using the tool named “Looper”, provided that the primary sequence of the protein is known.¹⁸² In our efforts to find novel inhibitors for SpdSyn, we have prepared the two mentioned crystal structures to use in our drug discovery efforts.

The PDB file (ID: 2PT9) was retrieved from the PDB. In the crystal structure, three protein chains have been co-crystallised in the asymmetric unit of 2PT9. Using the protein report tool for guidance, Chain A was selected for preparation as it had the least amount of missing and/or incomplete residues and the ethylene glycol crystallisation factor was removed. Firstly, the entire complex was typed with the CHARMM forcefield, with partial charges estimated by Momany-Rone. Following this, all basic amines were set with a +1 formal charge. (Subsequently, the tool which is able to calculate residue titration curves was found. The pH can then be set at a pH of 7.4 to automatically ionize the protein.) During the typing process, hydrogen atoms are added into the structure. Although most of these hydrogens are correctly placed on-the-fly by DS, in some instances additional calculation is required. As seen in Figure 4-2, hydrogens may be trapped in local energy minima while a more energetically favourable conformation may exist. To traverse the energy gradient, a standard dynamics cascade may be used.

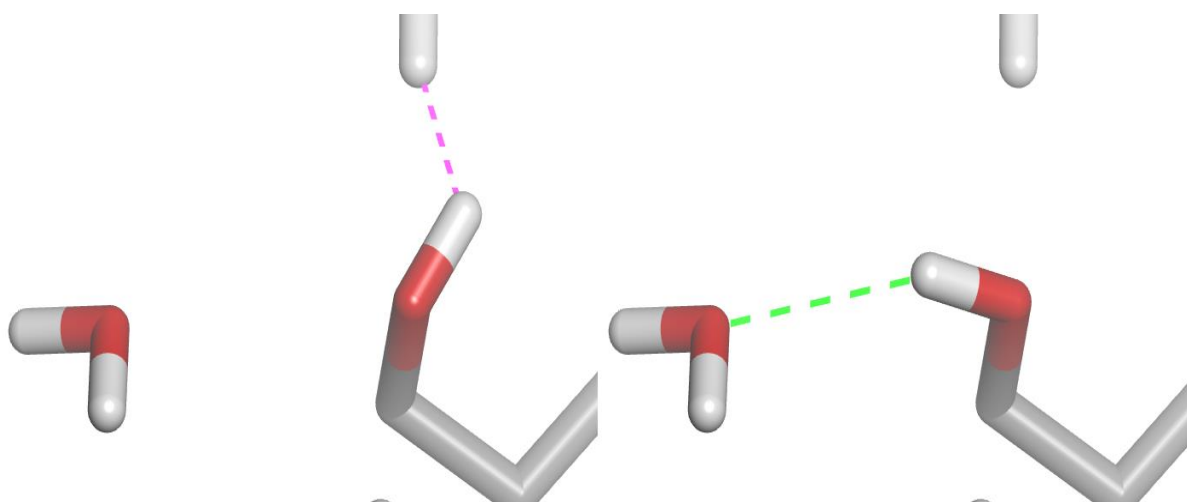


Figure 4-2 Optimisation required for hydrogen atoms.

Hydrogens may exist in various positions based on the parent heavy atoms. Often, simply performing minimisation would not lead to the desired state. Here, the hydroxyl's lowest energy state is when making a hydrogen bond to the adjacent water molecule. However, the two methyls (above and below) prevent this and requires additional energy to be entered into the system before the global minimum may be reached.

For the standard dynamics cascade, all atoms except hydrogens were placed under fixed atom constraints. This would keep all the heavy atoms (non-hydrogens) that were obtained from the PDB fixed in place, while only the hydrogen atoms go through the rigorous heating and equilibration cycles. Seen in Figure 4-3 is a graph of the total energy of the system as it progresses through the heating steps. The row indexes correspond to each heating step which, in turn, correlates to an alternative conformation set of the hydrogen atoms.

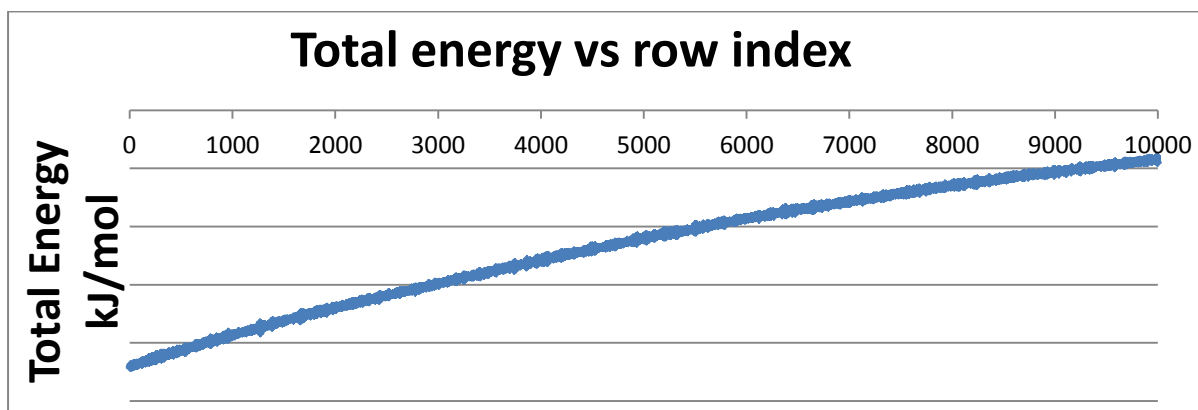


Figure 4-3 Increasing the total energy of the system by the increase of the temperature. This is accomplished by increasing the kinetic energy of the atoms.

After the heating phase, the simulation is allowed to equilibrate during which the kinetic energy becomes evenly dispersed throughout the system. This allows the potential energy wells to be searched. During this cascade, 10 000 data points were collected, as seen in Figure 4-4.

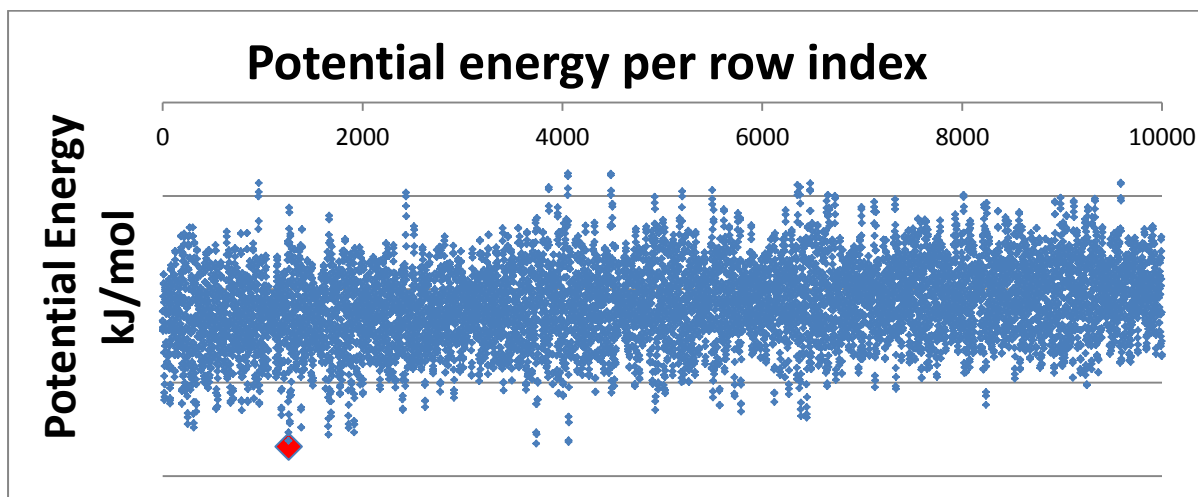


Figure 4-4 Kinetic equilibration of the introduced energy.

This allows an effective simulation of the atoms as they move around. During this period, snap shots are taken of their relative position and the potential energy components can be determined. This effectively allows the potential energy wells to be explored. Indicated in red, is the lowest potential energy point in this data set.

The conformation with the lowest potential energy was picked, but unfortunately was found to contain a hydrogen atom steric clash near the active site. To once again expel this steric

clash which is trapped in a local minimum, this protein conformation was then resubjected to another standard dynamics cascade (see figure Figure 4-5) , which then further reduced the potential energy from -3226.09 kJ/mol to -3255.05 kJ/mol without any steric clashes.

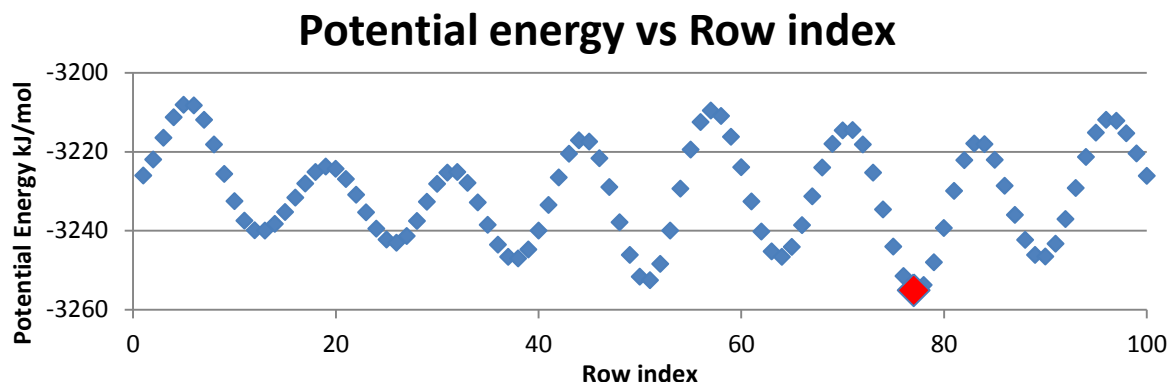


Figure 4-5 Graph of the potential energy as the production run of the standard dynamics cascade proceeds.

The red data point represents the point of lowest energy which was carried through to the next steps.

Next, histidine tautomeric states were adjusted manually and, if required, nearby water molecules were rotated to optimise hydrogen bonding networks. Seen in Figure 4-6 is a typical example of such an occurrence where manual manipulation is still required for optimisation. In this case, three alterations needed to be made, which included: histidine tautomeric state selection, rotation of the phenolic hydrogen and rotation of the adjacent water molecule.

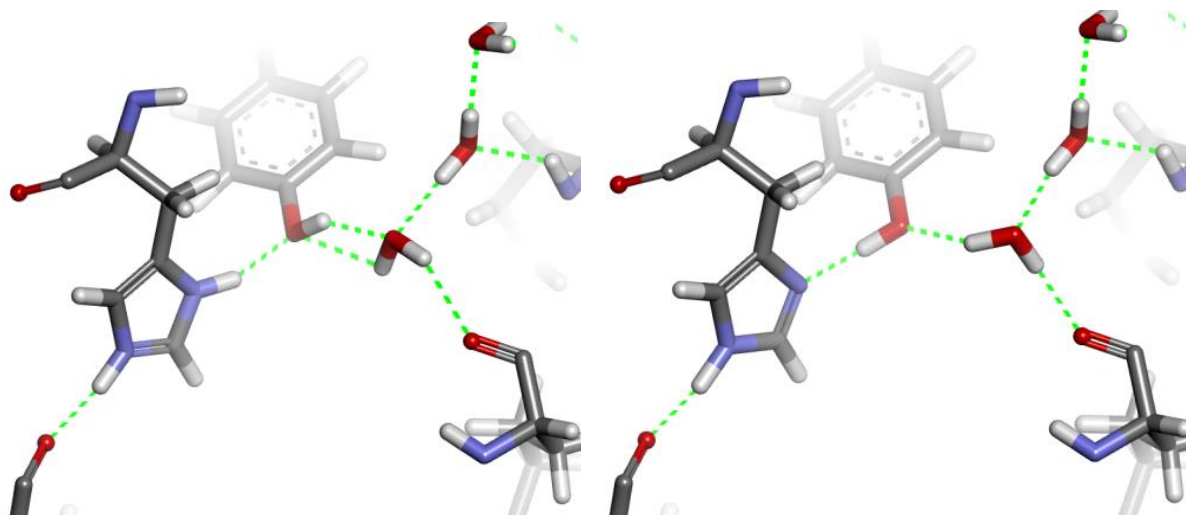


Figure 4-6 Example of histidine charge / tautomeric optimisation.

After a cascade, even though most minima have been obtained, manual intervention may be required in some cases. Here, it was evident that the automatic pKa calculation placed a charge on this Histidine residue. This was corrected by the removal of the charge, choosing the correct tautomeric state, rotation the adjacent hydroxyl bond and rotation of the entire adjacent water molecule yielded the correct state. Although the water molecule was rotated from its original position, this will be corrected for in the final minimisation step.

Following this, the protein was minimised for a final time using much more rigorous parameters. Only the peptide backbone was constrained and an explicit solvent model was used to increase the accuracy of the simulation. This step would alleviate most of the crystal packing energy and attempt to simulate the protein in the solvated state. The final minimisation criteria for this minimisation was $\text{RMS} = 0.1 \text{ \AA}$. To test whether the prepared protein is adequate for modelling purpose, the co-crystallised ligand was then re-docked back into the active site. In this case, 4MCHA was successfully redocked, back into the structure with an RMSD of 0.508 \AA compared to the original crystal pose (seen in Figure 4-7). The prepared protein structure was now ready for use.

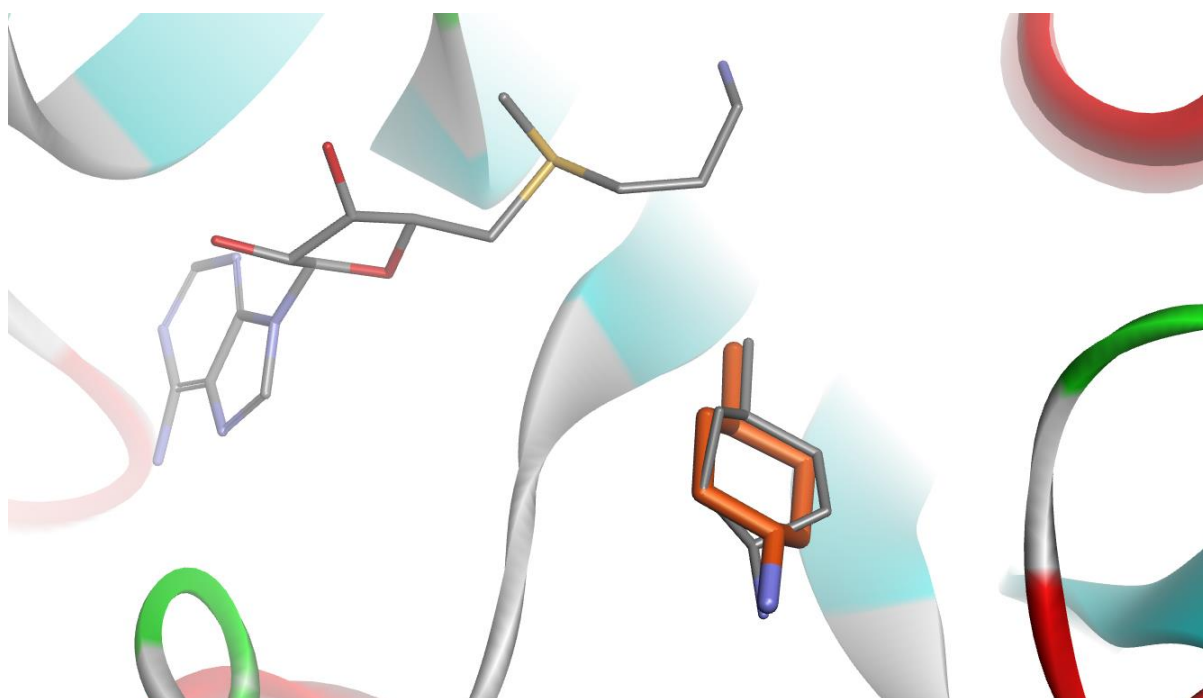


Figure 4-7 Successful redock of 4MCHA (in orange) back into 2PT9.

This gave indication that the prepared protein was ready to use for further modelling purposes. For reference, the original crystal ligands are also shown (grey carbons).

One of the first applications for this newly prepared model would be to use it to perform vHTS in efforts to quickly identify possible hits.

Virtual high-throughput screens

Virtual high-throughput screens (vHTS) require significant computational time to scan large databases of candidate molecules. To reduce this cost, vHTSs are performed so as to eliminate unsuitable candidates as soon as possible while using minimal computational resources. After each step, increased computational resources are applied to examine candidate molecules in greater detail, to culminate in the most rigorous simulation only at the final stage of the vHTS. Using this strategy, expensive computational time is not wasted on candidate molecules which could have been eliminated using simple analyses.

For this dissertation, two types of vHTS were performed: a pharmacophore based screen and a hotspot based screen.

Pharmacophore based screen

Our first vHTS was performed on our newly prepared model of pfSpdSyn (PDB ID: 2PT9). To ensure that the hits were easily obtainable, we used the Sigma Aldrich catalogue as a database of compounds because any hits coming from this exercise could then easily be purchased.

Database preparation

To start our vHTS, we obtained a previously prepared SD file containing the entire Sigma-Aldrich vendor catalogue. We decided to generate conformations before-hand, which increased the 13 573 compounds available in the SD file to 217 701 compounds and their corresponding conformations using the build 3D database tool. Following this, the pharmacophore had to be constructed.

Pharmacophore construction and screening

Construction of the pharmacophore is a critical part of the vHTS and requires careful consideration throughout this phase, as it directly influences how the possible hit interacts with the protein. Firstly, because of the significant energy penalty involved in the displacement of conserved, structural water molecules, they needed to be identified. This was done by superimposing the apo enzyme (PDB ID: 2PSS) onto the 4MCHA complex protein (see Figure 4-8). From this, possible hydrogen bond donating and accepting interaction sites were added (see Figure 4-9). Next the active site shape of 2PT9 was determined and defined using Q-site,¹⁸³ which would impose spatial constraints on the search. Following this, a hydrophobe was placed in the middle of the active site and a pi-interaction was placed coming off a tyrosine residue.

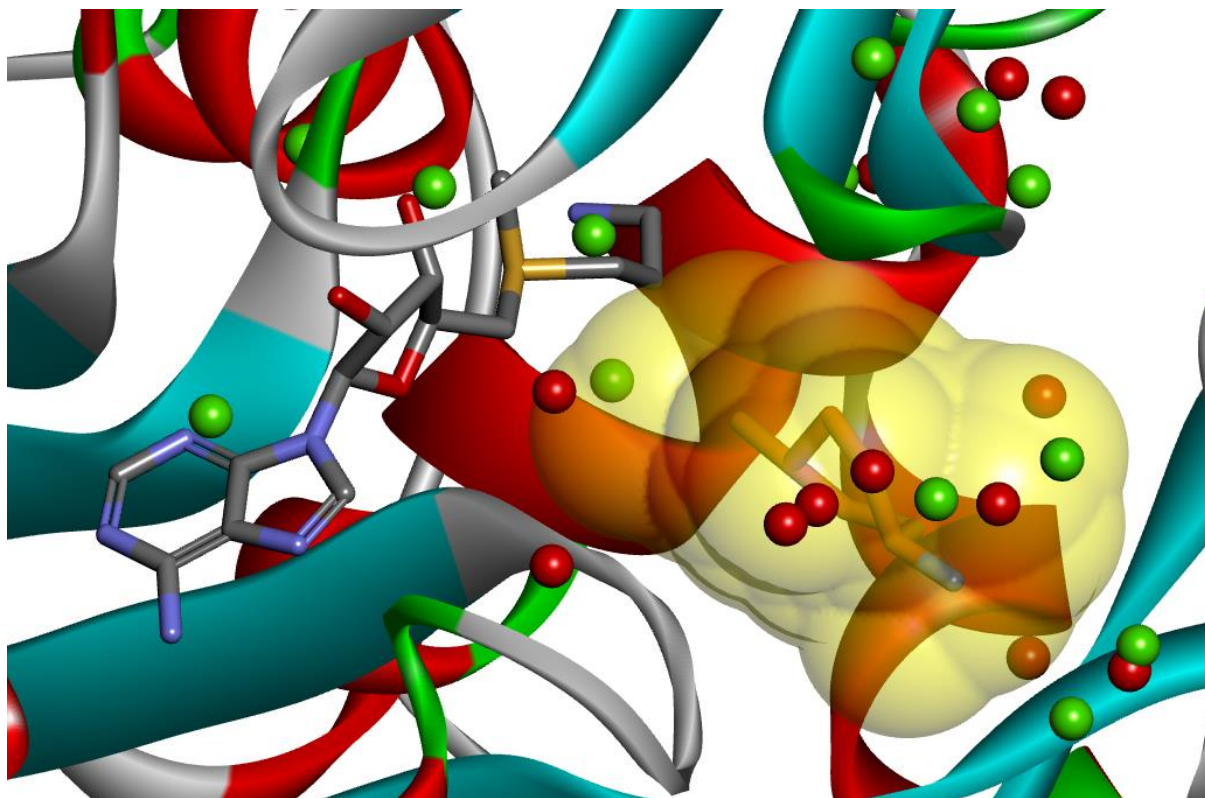


Figure 4-8 Development of the pharmacophore inspired by 4MCHA (carbons in orange). The co-crystallised water molecules are shown in red and in green are the water molecules from the apo enzyme. Although there are discrepancies, it is still apparent which molecules are conserved and should be retained in the active site. The yellow region defines the shape of the active site.

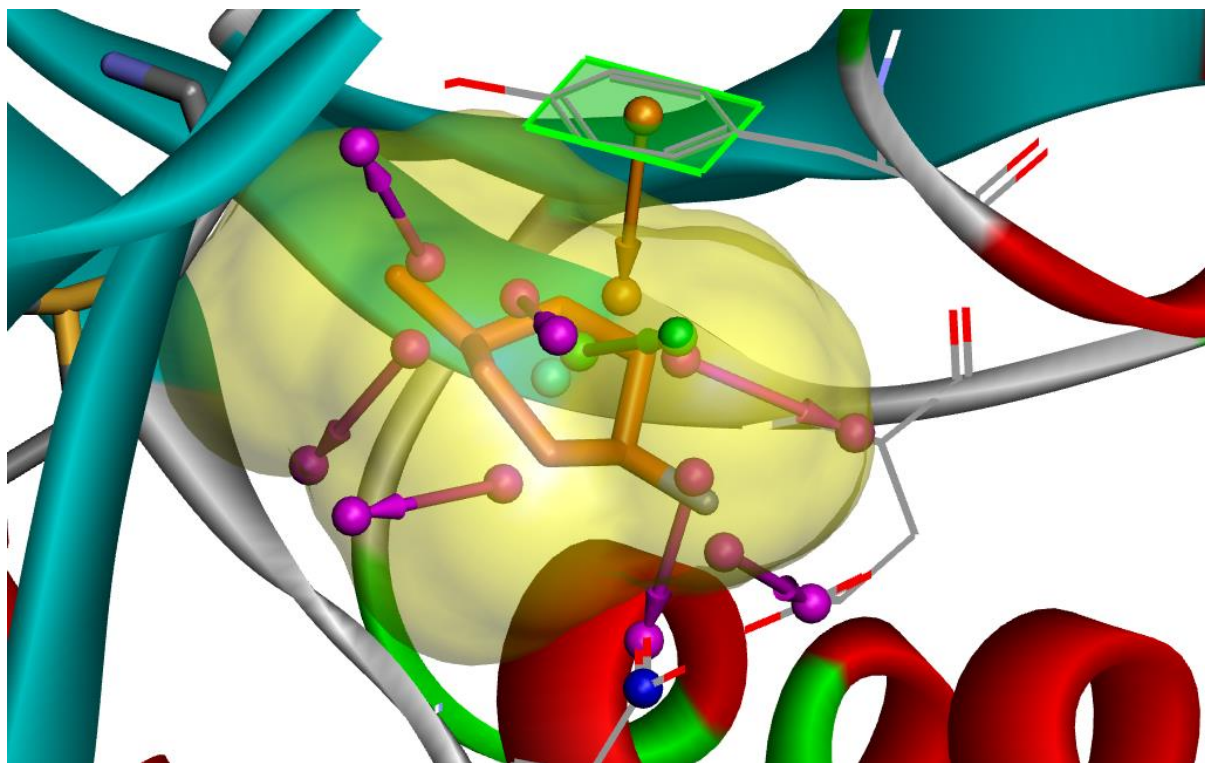
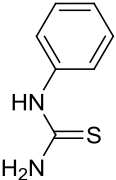
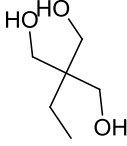
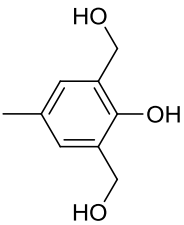
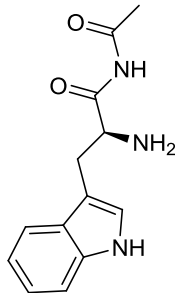
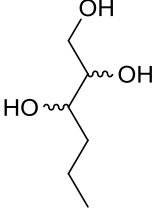


Figure 4-9 Completed pharmacophore inspired by 4MCHA (carbons in orange). Shaded area indicates the shape of the active site; purple arrows indicate hydrogen bond donors; turquoise dot in the middle indicates hydrophobe; orange arrow indicates a pi-interaction; blue dot indicates a negatively charged amino acid.

Chapter 4 – Molecular modelling (Spermidine Synthase)

The SD file containing the Sigma compounds and their various conformations was then screened against our pharmacophore. This resulted in 273 hits and after removing duplicates, 25 unique molecules were obtained. Having eliminated compounds from the database which would not fit the requirements of our pharmacophore, we were now in a position to apply far more rigorous and computationally expensive techniques to the remaining 25 compounds. For this, we turned to CDOCKER for docking. The results with the 5 highest CDOCKER scores were selected for purchase and were subsequently tested at a concentration of 100 μ M in an enzymatic assay. The results of this vHTS are presented in Table 4-1.

Table 4-1 Results of the Sigma-Aldrich catalogue screen by pharmacophoric features.

	Control					
		1	2	3	4	5
CDOCKER interaction energy	47	30	28	8.2	-24	-31
CDOCKER score	40	14	12	-4.2	-82	-131
Enzymatic assay (100 μ M inhibition)	52%	7%	Negligible	Negligible	Negligible	Negligible

The difference between the CDOCKER interaction energy and the CDOCKER score is that the former only accounts for interactions that the ligand has with the protein. The score includes the interaction energy, but also accounts for the conformational strain which the ligand is subjected to in that proposed bioactive conformation.

Unfortunately, only one compound showed activity in the enzymatic assay. Also, barring the limited data points available, we can also see that there may be a trend of activity which correlate the CDOCKER score and the enzymatic activity.

2I7C

Although the inhibitor 4MCHA is potent, it only has interaction with a small region of the active site of SpdSyn. To expand the number of possible interactions available for new inhibitors, we prepared the structure with PDB ID: 2I7C. This structure is co-crystallised with the other mentioned potent inhibitor, AdoDATO. Compared to 4MCHA, AdoDATO is approximately four times larger and has interaction with the entire putative binding site. This means that the protein is in a conformation which is slightly more accommodative towards both AdoDATO and other molecules of similar sizes. For this reason, we used this protein structure for the remainder of this dissertation. This included a second vHTS (later in this chapter) and the design and docking of novel antimalarials (Chapters 5 and 6).

Protein preparation

Similarly to PDB ID: 2PT9,¹²² three proteins were crystallised in the asymmetric unit of PDB ID: 2I7C and also similarly to PDB ID: 2PT9, chains B & C made up the homodimer. Although in these two chains, several residues had either incomplete or alternate conformations, in this case we deemed it to be more appropriate to use the entire homodimer to prepare the protein structure. Again using the “Clean protein” tool, these residues were corrected and the protein was then typed with the CHARMM forcefield. Ligand ionisation states were manually assigned while for the rest of the protein the residue states were calculated using the “Calculate Protein Ionization and Residue pKa” tool. The histidine tautomeric states were again manually adjusted. Following this, the standard dynamics cascade was performed with all non-hydrogen atoms fixed in place. Once again, the lowest potential energy conformation was selected and monitored for bumps.

The final minimisation was performed with the explicit solvent model, incorporating Particle Mesh Ewald electrostatics. The crystal waters were kept in place. To test the model, CDOCKER was successfully used to dock the co-crystallised ligand, AdoDATO with an RMSD of 0.5369 Å, although this included a slight conformation mismatch in the region of the sulphur atom (see Figure 4-10).

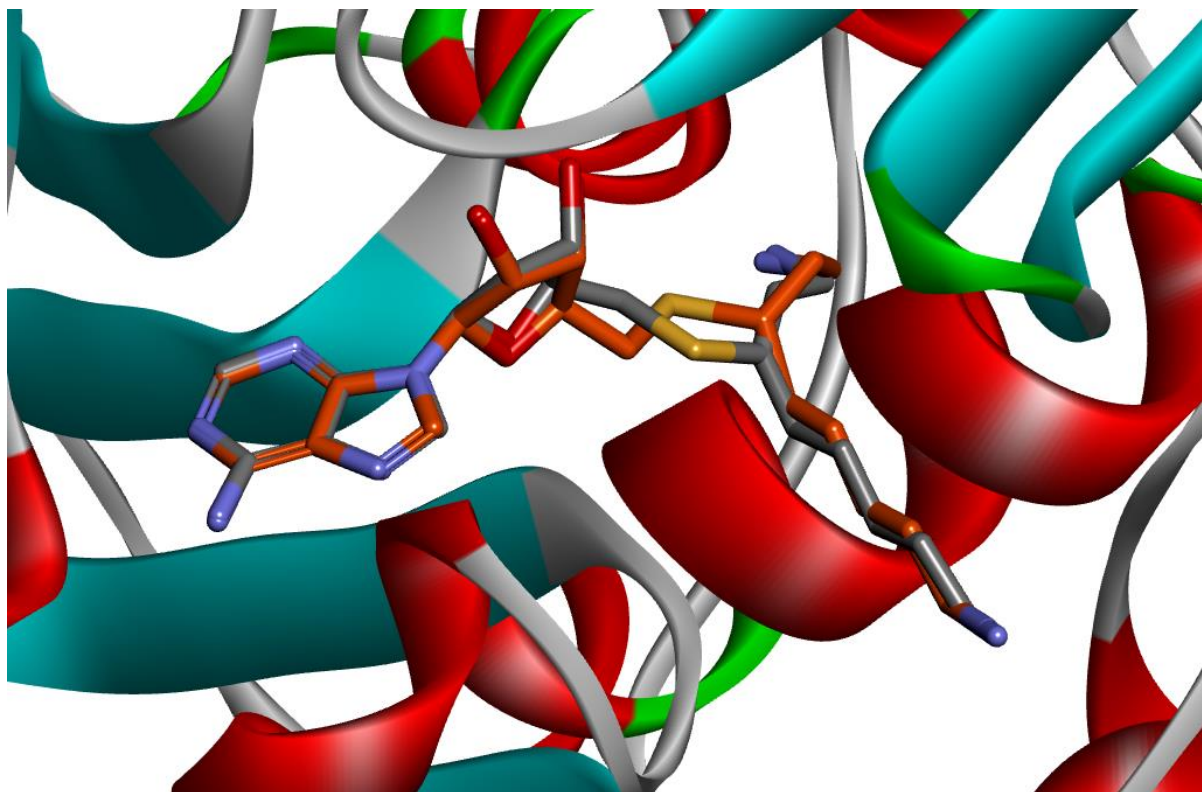


Figure 4-10 Prepared crystal structure of 2I7C.

The co-crystallised ligand, AdoDATO, has successfully been redocked (orange carbons). Superimposed for reference is the ligand from the actual crystal structure (carbons in grey). Only mild deviation is evident in the linker sulphur.

Hotspot guided screen - LibDock

In another attempt of vHTS, a significantly larger database was used in combination with the LibDock¹⁸⁴ algorithm.

Database preparation tool

DS includes numerous tools to prepare a chemical database for vHTS and this process can be streamlined using Pipeline Pilot (PP). For this dissertation, a custom database preparation protocol was written to automate the preparation of chemical databases and was subsequently used to prepare such a database from the ChemDB database.¹⁶⁹ This reduces user input requirement, reducing the chance of user error while also providing better overview of the entire process. Seen in Figure 4-11 is a representative flow diagram of the workflow. Subroutines that are clustered together are expanded below the grouped protocols.

In the first block the protocol starts by reading the input file(s) which store the molecule entries (typically in SD format). The user needs to specify how many entries are read at a time. This has two purposes: If too many files are passed into the pipeline at a single time, the system RAM could be overloaded and a crash would result. Also, the output files are written as multiple smaller files which are more manageable rather than a single huge file.

Chapter 4 – Molecular modelling (Spermidine Synthase)

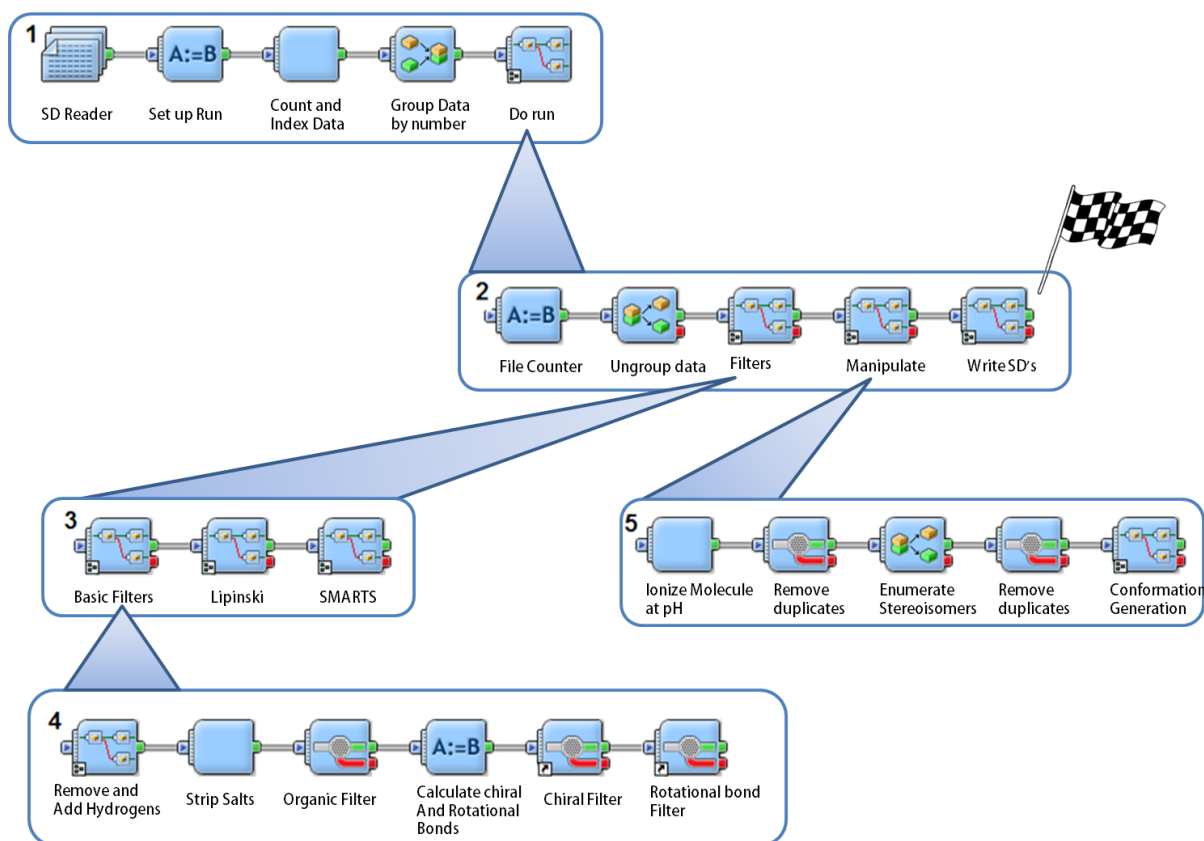


Figure 4-11 Schematic of Pipeline Pilot workflow to prepare large databases for screening.

In the second block, the entries which are now in the pipeline are passed along to the various stages of preparation. Starting with the computationally least expensive processes, compounds which fail the criteria are removed in the filters section: blocks 3 and 4. Firstly, some databases come with hydrogens attached to the compounds, whereas others do not. By resetting this state manually, the protocol ensures that hydrogens are added. Next, if metal salts are present, these are stripped from the molecules. Then, the organic filter removes any molecule which does not entirely consist of C, H, O, N, S, P, F, Cl, Br and I. The protocol then calculates two properties of each molecule: the number of chiral centres as well as the number of rotational bonds so that the user can specify the removal of them if need be. Molecules with chiral centres are difficult to synthesise and molecules with ≥ 10 rotational bonds are difficult to handle *in silico*. The user may then specify whether molecules which exceed criteria in this regard should be discarded. Following this, Lipinski (described earlier) and SMARTS filters are applied as specified by the user. SMARTS filters are a variety of chemical signatures which may describe a variety of structural features of molecules which may be discarded. These range from chemically reactive moieties, known toxicophores to synthetically challenging moieties. The complete list is provided below.

Table 4-2 SMARTS filters

Filters applied in database preparation

Sulfonyl halide	Epoxides or Thioepoxides or Aziridines
Acid anhydrides	Nitros
Acid halides	Quaternary C-Cl-I-P-S
HOBt esters	Peroxides
Phosphoramides	Lawesson reagent and derivatives
Di or Triphosphates	Triflates
Crown ethers	Aromatic azides

Filters not applied in database preparation

Disulfides	Thiocyanates
Thiols	Carbodiimides
Fluorines	Sulfonates
Cl-Br-I	Cyanidin derivatives
Cyanohydrines	Thiocyanate
Sulfate esters	Benzylic quaternary Nitrogen
Pentafluorophenyl esters	Phosphoramides
Paranitrophenyl esters	Beta carbonyl quaternary Nitrogen
Beta carbonyl quaternary N	Acyhydrazide
Acyl cyanides	Phosphoranes
Sulfonyl cyanides	Nitroso
Thioesters	Isonitrile
Cyanamides	Triacyloximes
Cyano	Cyanohydrins
Isocyanates or Isothiocyanates	Acyl cyanides
Paranitrophenyl esters	Sulfonyl cyanides
Isocyanates and Isothiocyanates	Cyanophosphonates
Polyenes	Azocyanamides
Carbazides	Azoalkanals
Squalestatin derivatives	Aliphatic methylene chains 7 or more long
Sulfonate ester	Compounds with 4 or more acidic groups
Phosphonate ester	Four numbered Lactones
Isonitriles	Di and Triphosphates
P or S Halides	Betalactams
Carbodiimide	Quinones
Aliphatic alcohols	Cytochalasin derivatives
2-4-5 trihydroxyphenyl	Saponin derivatives
2-3-4 trihydroxyphenyl	Cycloheximide derivatives
Hydrazothiurea	Monensin derivatives
Sulfonates	Alkyl halides

Once the filtration steps have been completed in blocks 3 and 4, the pipeline is followed in block 2. This leads to the manipulation step, which is expanded in block 5. The molecules are then ionised at physiological pH of 7.4, followed by the enumeration of the stereoisomers. After each of these two steps, duplicate molecules are prone to occur and are then filtered out. Finally, conformations are generated from the entries by Catalyst. Conformation generation is a computationally expensive process and will drastically increase the time required to process the protocol. Also, the storage space required for the output file is also increased significantly. It is up to the user to decide whether to generate

conformations beforehand or calculate them on-the-fly during the database screening phase. Finally, the output entries are saved as SD files for storage as a prepared database. The ChemDB database was prepared using this protocol, with the parameters listed in Table 4-3. This reduced the 7.1 million unique compounds to 3.2 million compounds, which included pre-generated conformers.

Table 4-3 Database preparation parameters

Process	Parameter	Justification
Chiral filter	≤ 2	Synthetically challenging
Rotatable bonds filter	≤ 10	Difficulty during docking
SMARTS filters	Underlined in list	Various

Scoring functions

To determine which scoring functions would be best suited towards a particular system, a positive control is required. This entails using poses of known inhibitors in the active site and applying scoring functions to those poses. Functions which score known inhibitors high can then be used to score potential new inhibitors. To find appropriate scoring functions for our system, we used results from Jacobsson *et al* as positive control.¹⁸⁵ In their study, they identified 28 compounds in a vHTS, of which 4 showed notable enzymatic activity. These 28 compounds were docked into the active site of the protein using CDOCKER. All available scoring functions were applied to the resultant poses and the scores were normalised according to the maximum score obtained for each scoring system. This data was then plotted and visually examined to compare the scores (see Figure 4-12). LigScore1 returned very good (>85) results for 3/4 inhibitors. PLP2 also scored 2/4 inhibitors well. LigScore2 rates almost all the inhibitors very high and would therefore not be appropriate. LigScore1 and PLP2 would be used in our vHTS scoring step. The magnitudes of these two scores were also recorded to be able to directly compare these scores with the results which we may obtain (see Table 4-4).

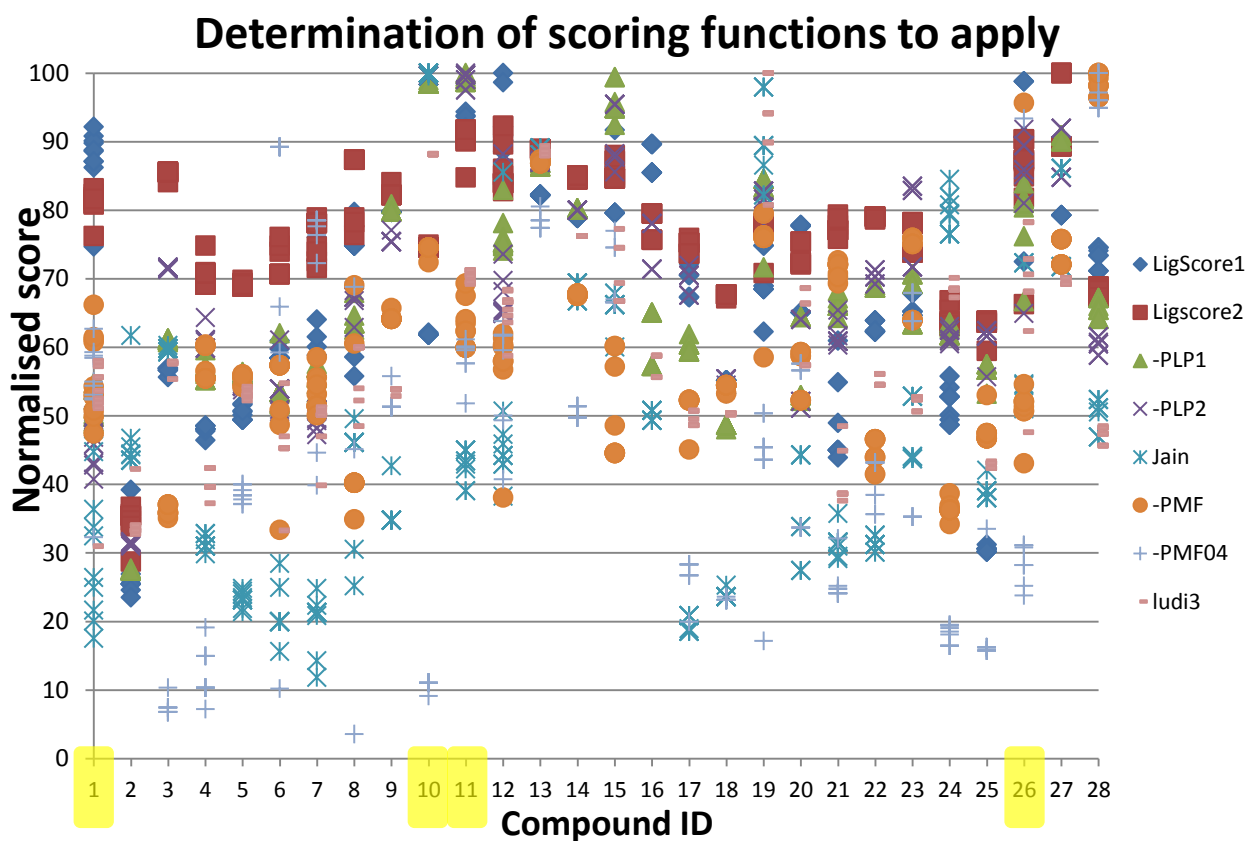


Figure 4-12 Different scoring functions applied to the various potential inhibitors reported by Jacobsson *et al.*¹⁸⁵

The 4 compounds highlighted in yellow are known to be good binders to SpdSyn and were used as positive controls.

Table 4-4 Range of magnitudes of the scores which delivered potent results.

Compound	LigScore1 range	-PLP2 range
1	5.1 - 6.4	50 – 60
10	4.2	120
11	6.1 - 6.4	117 – 121
26	4.9 - 6.7	78 – 111

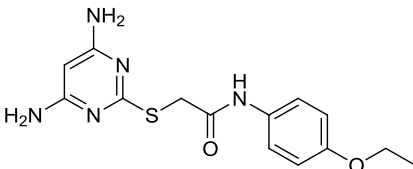
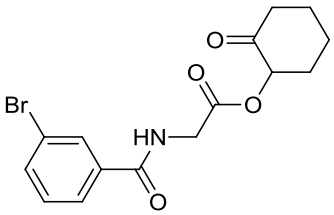
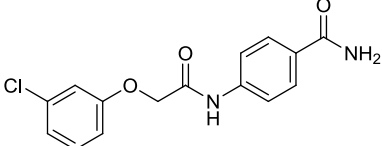
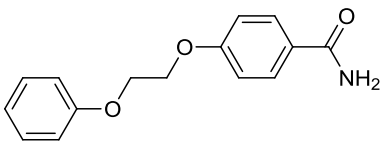
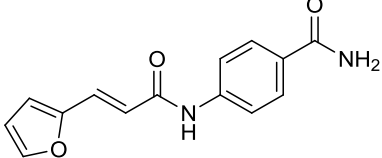
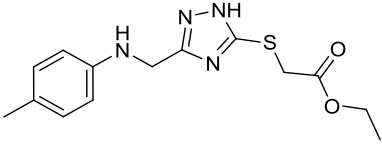
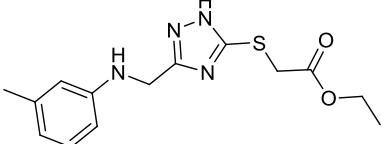
Libdock sequence

The LibDock sequence was then used to screen this prepared ChemDB database of 3,277,039 possible ligands. Of these, 2,636,460, either failed to align to the hotspots or had mismatching polar/non-polar regions. The remainder resulted in 3,945,513 successful poses. Appropriate scoring functions were then needed to filter down the results.

After applying consensus scoring on the top 10%, 4 million poses were reduced to 13 967 poses. After removing duplicates, aqueous solubility and intestinal absorption criteria, 2470 unique compounds remained.^{43, 186} Using TOPKAT AMES 3.1 mutagenicity prediction, all compounds which had more than 10% chance to have toxicity were removed, resulting in 1 661 unique compounds. Of these compounds, 943 were successfully docked in the protein structure using CDOCKER. These poses were then visually examined, of which 7 were deemed to be possible hit compounds. Furthermore, the poses returned from CDOCKER were compared to those from LibDock to assess whether the same pose was returned. Also, the LigScore1 and PLP2 scores were reapplied for these new poses. The results are presented in Table 4-5.

Chapter 4 – Molecular modelling (Spermidine Synthase)

Table 4-5 Possible hits of the vHTS using LibDock from 3.2 million compounds from ChemDB database

#	ID	Structure	-CDOCKER Score	LigScore1 Score	-PLP2 Score
1	3842842		44	5.2	100
2	5716589		43	4.5	101
3	4144030		40	5.6	103
4	7499777		35	4.5	91
5	4409360		32	3.9	85
6	3637839		25	5.4	95
7	3638125		21	5.4	94

Given that CDOCKER returned the same pose as LibDock for 5 of the 7 compounds, relatively high –CDOCKER scores and that both LigScore1 and –PLP2 scores are within the same range as the positive controls used from Jacobsson *et al*,¹⁸⁵ it is very likely that at least 1 or more of these compounds could have very notable inhibition towards SpdSyn activity. Unfortunately, we were unable to test these compounds for activity to classify them as “hits”.

However, should one of them prove to be effective, there is a good likelihood that those compounds would advance several more steps during the hit to lead phase, bearing in mind that these 9 compounds have already passed through numerous filters (Lipinski, toxicity, solubility, etc.) to reach this stage.

To the laboratory...

This point also highlights the limitations of computational methods, which are based on theory. These methods would almost always deliver results, but they remain just that: theoretical. Only once actual, *physical* compounds are tested, can the results of these methods be proven. Then, if these compounds show activity, comprehensive and systematic structure activity relationship studies (SAR), with frequent testing, have to be performed on these molecules to thoroughly explore the chemical space of the active site. These SAR studies involve installing various groups on various points on the molecule to determine what affects these may have on the activity. Some of these groups include: electron withdrawing groups, electron donating groups and spacers.

Although modelling may guide the drug discovery process at this point, the bulk of this work still resides within the synthetic organic chemistry laboratory. Only with real molecules that are actually tested can reliable data points be sourced to monitor the development of a possible hit or lead. However, the only way how the proposed inhibitors can be obtained, is using synthetic techniques.

5. Spermidine Synthase – Electrophilic adenosine

Design

Our synthetic efforts would stem from the transition state analogue AdoDATO (refer to Figure 5-1). AdoDATO is a conjunction of the two natural substrates of SpdSyn: dcAdoMet & putrescine. It has hydrogen bonding interactions to SpdSyn via the nucleoside and the two primary amines, including a pi-stacking interaction (refer to Figure 5-2).

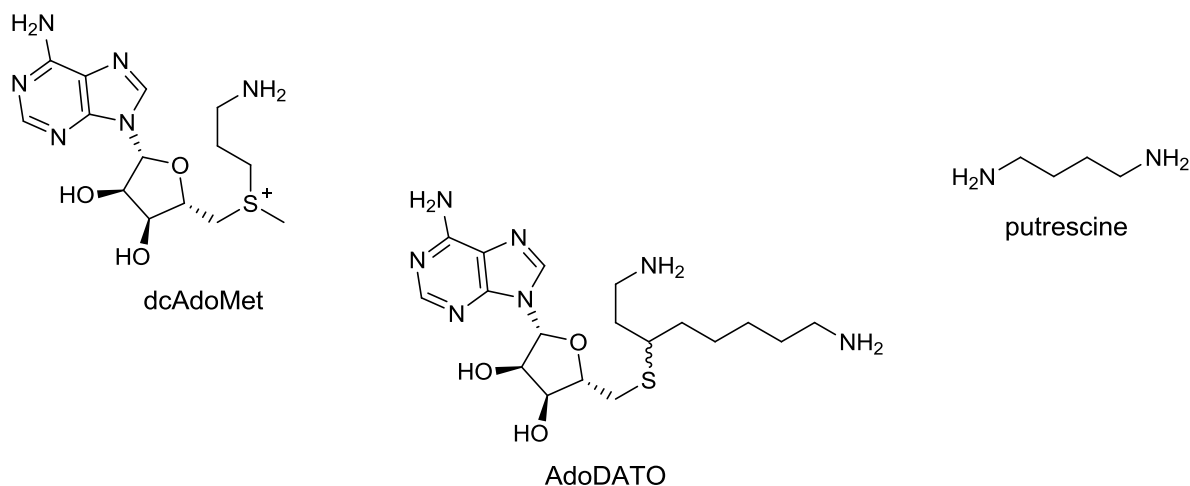


Figure 5-1 Natural substrates of SpdSyn
AdoDATO is a transition state analogue of the two natural substrates of SpdSyn; dcAdoMet & putrescine.

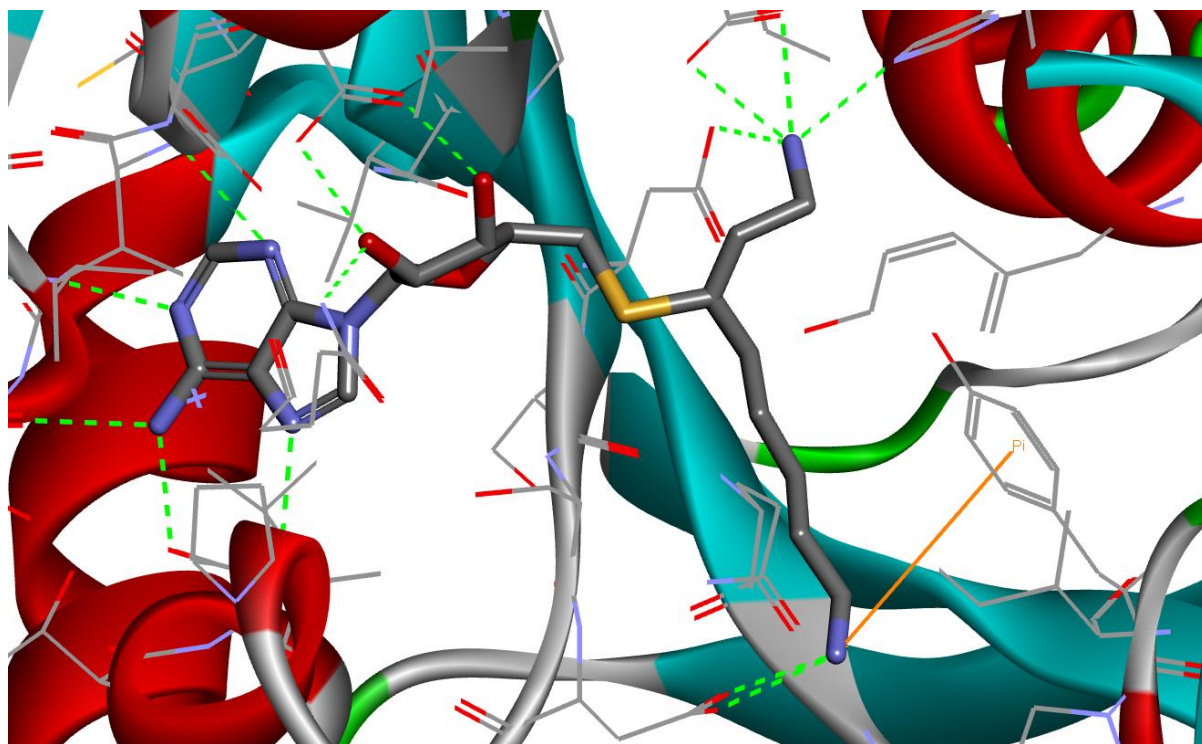
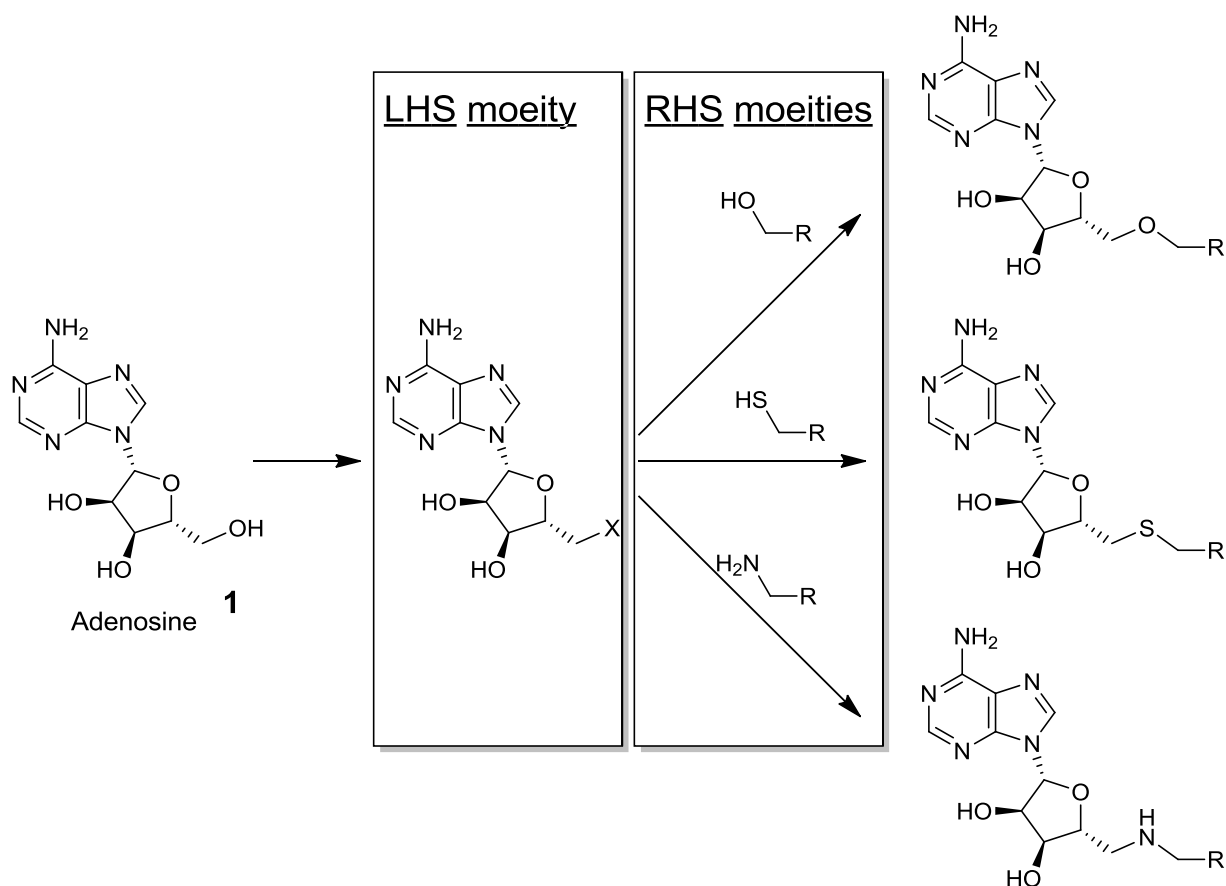


Figure 5-2 AdoDATO interactions areas with SpdSyn.
Including the nucleoside, dcAdoMet has interactions with the two cationic pockets (to the upper and lower left) and the hydrophobic region between them.

Chapter 5 – Spermidine Synthase (Electrophilic adenosine)

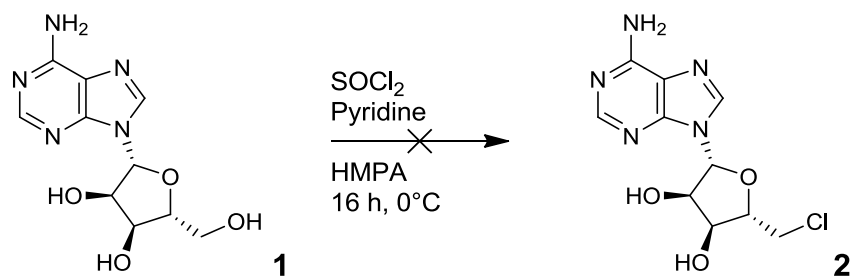
To begin our library (refer to Scheme 5-1), we chose adenosine **1** to act as a starting scaffold. It is readily available and has a natural affinity for our target. Importantly, from a synthetic point of view, the primary hydroxyl offers a versatile handle for functionalisation. After the nucleoside has been prepared, we envisaged to create a library of compounds which may act as nucleophilic right hand side (RHS) moieties, similar to the alkyl/amino functionality found in AdoDATO and putrescine to allow linking to the adenosyl on the left hand side (LHS).

Scheme 5-1 Synthetic strategy to prepare SpdSyn inhibitors



Synthesis – Chloroadenosine

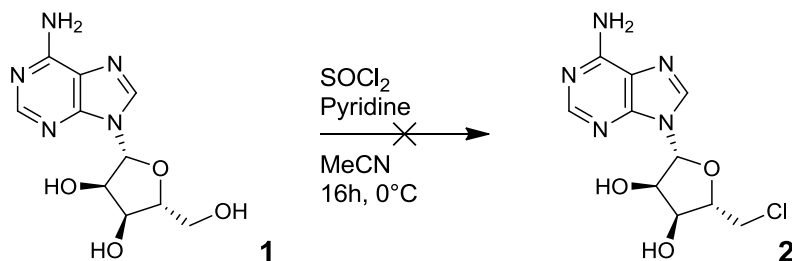
Scheme 5-2 Attempted synthesis of 2



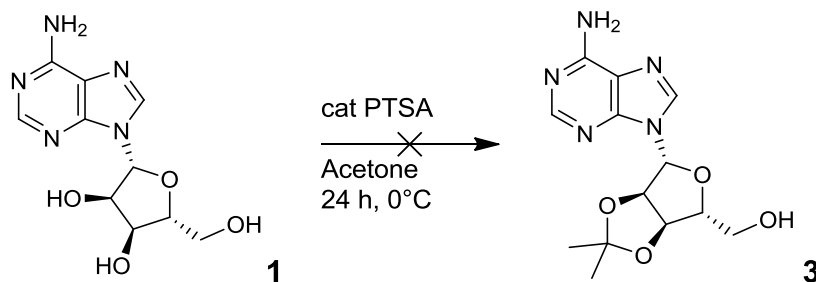
To prepare the LHS as an electrophile, the synthesis began by the conversion of the primary hydroxyl into a chloride using thionyl chloride in hexamethylphosphoramide (HMPA).¹⁸⁷ To neutralise the HCl by-product, pyridine was added. Although the secondary hydroxyls would also be susceptible to conversion, their reduced reactivity would offer the required chemoselectivity.

Unfortunately, no new products were evident by TLC after an extended period. After this failure, the use of HMPA as solvent was abandoned for three major reasons: it is both carcinogenic and toxic; it is difficult to remove from the organics and also requires vacuum distillation prior to use.

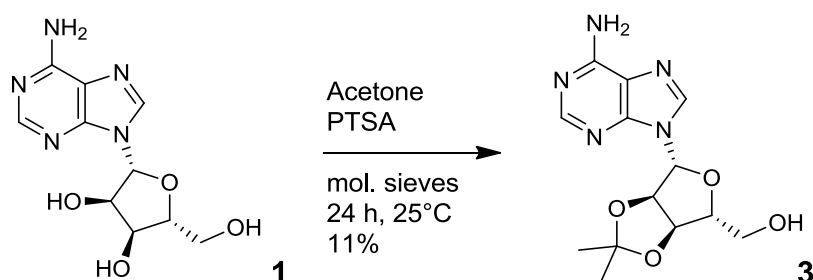
Scheme 5-3 Alternative attempted synthesis of 2



The chlorination was re-attempted in acetonitrile using a procedure developed by Robins *et al.*^{188, 189} The reaction proceeded as expected, up to the point that the excess thionyl chloride was neutralised with a mixture of MeOH and aqueous NH₄OH. As this was performed, the neutralisation was monitored by TLC and it was observed that the newly formed product reverted back to starting material. We suspected that the basic aqueous solution was nucleophilic enough to displace the newly created chloride.

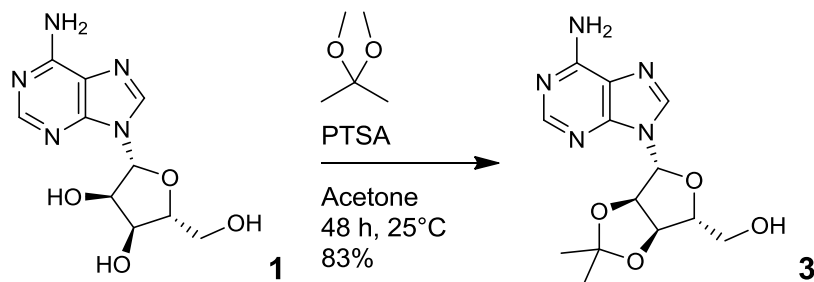
Scheme 5-4 Attempted synthesis of 3

On the suspicion that it might be easier to handle the compound if the polarity was reduced and to avoid potential future chemoselectivity issues, it was decided to protect the vicinal diol as a ketal. Using a catalytic amount of para-toluene sulfonic acid (PTSA), the ketalation was attempted in acetone. Owing to Le Chatelier's principle, the large excess of acetone would drive the reaction forward, mitigating any water present in the technical grade acetone. Unfortunately, these conditions did not produce any new products.

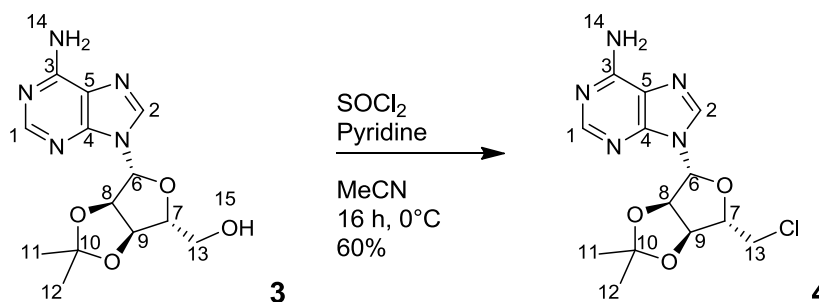
Scheme 5-5 Synthesis of 3

Having been unsuccessful in our first attempt to install the ketal protecting group, we now reattempted the reaction with modifications.¹⁹⁰ Firstly, strict anhydrous conditions were used to remove the water by-product which would push the reaction equilibrium backwards. This included the addition of molecular sieves to the reaction vessel, thorough purging of the atmosphere in the vessel with anhydrous nitrogen, hot assembly of the glassware from the oven pre-drying of all reagents and solvents. Also, even though only a catalytic amount of H⁺ is required theoretically, 10 equivalents of PTSA catalyst were used to expedite the reaction.

The successful synthesis of the ketal was confirmed using ¹H NMR spectroscopy by the presence of two upfield singlets which both integrated for 3 protons. These signals corresponded to the two methyls from the acetone molecule. It is noteworthy that these two methyls appear as two distinct signals, as they are not chemically equivalent when considering the three dimensional structure of the chiral molecule.

Scheme 5-6 Alternative synthesis 3

In an alternative method to produce the ketal, the irreversible reaction using dimethoxy propane was investigated.¹⁹¹ Four equivalents of DMP were added and the reaction was performed in acetone solvent. Using 12 equivalents of PTSA, the reaction went to completion in 48 hours. Again, the successful synthesis was confirmed using ¹H NMR spectroscopy. The obtained spectrum matches the molecule obtained by the synthetic procedure using only acetone.

Scheme 5-7 Synthesis of 4

Once the protected adenosine was isolated and purified, the chlorination was performed again using the same conditions as in Scheme 5-3.¹⁸⁹

In tandem with a change in the chromatographic retention factor, ¹H NMR spectroscopy confirmed that the chloro-adenosine was obtained by the disappearance of the hydroxyl signal, a slight change in the chemical shift signals and the complete assignment of all signals (refer to Table 5-1).

Table 5-1 ^1H NMR spectroscopic signals of **3** and **4**

Atom	3	4
1 and 2	8.34 (s, 1H)	8.33 (s, 1H)
	8.15 (s, 1H)	8.18 (s, 1H)
14	7.34 (s, 2H)	7.37 (s, 2H)
6	6.12 (d, $J = 3.1$ Hz, 1H)	6.23 (d, $J = 2.4$ Hz, 1H)
8	5.34 (dd, $J = 6.2, 3.1$ Hz, 1H)	5.48 (dd, $J = 6.3, 2.4$ Hz, 1H)
15	5.24 (t, $J = 5.6$ Hz, 1H)	N/A
9	4.96 (dd, $J = 6.2, 2.5$ Hz, 1H)	5.07 (dd, $J = 6.3, 2.9$ Hz, 1H)
7	4.21 (td, $J = 4.8, 2.5$ Hz, 1H)	4.34 (td, $J = 6.6, 2.9$ Hz, 1H)
13a	3.63 – 3.45 (m, 2H)	3.86 (dd, $J = 11.2, 7.0$ Hz, 1H)
13b		3.75 (dd, $J = 11.2, 6.2$ Hz, 1H)
11 and 12	1.54 (s, 3H)	1.54 (s, 3H)
	1.32 (s, 3H)	1.34 (s, 3H)

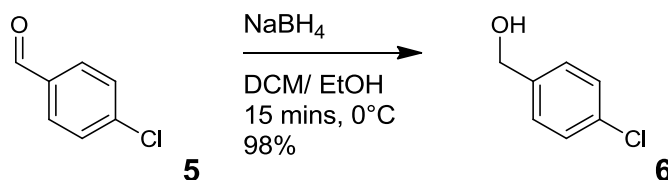
It was interesting to note that the methylene protons, situated at position 13, presented as 2 unique signals owing to the adjacent chiral centre (position 7). These two protons coupled very strongly to each other (11.2 Hz) and also to 7 (7.0 and 6.2 Hz). Furthermore, the coupling experienced at 7 is an average of the couplings from 13a and 13b.

Adenosine **1** was successfully made electrophilic at the required position using thionyl chloride. The task was now to expand the library to include a complementing suite of nucleophiles. These would comprise of easily deprotonated heteroatoms: hydroxyls, amines and thiols.

Synthesis – RHS moieties

Reduction

Scheme 5-8 Synthesis of 6



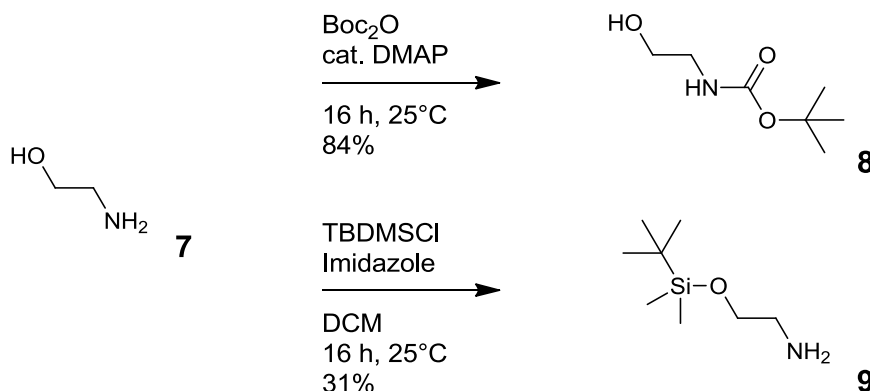
NaBH₄ is a mild reducing agent, typically used for the reduction of more reactive moieties like aldehydes or ketones. In this instance, it was used to synthesise the alcohol **6** from 4-chlorobenzaldehyde **5** in excellent yield.

¹H NMR spectroscopy confirmed that the reduction occurred by the disappearance of the aldehyde singlet and appearance of a broad singlet, integrating for 1 and another singlet, integrating for 2 protons. Additional Note however that it was not expected for the hydroxyl proton to couple to the methylene carbon protons. Additional characterisation was not performed as the compound has been synthesised previously.¹⁹²

This alcohol was planned to be the first molecule to be bonded to the chloro-adenosine **4**.

Chemoselective protection

Scheme 5-9 Synthesis of 8 and 9



Two more substrates to the chloro-adenosine could be made from ethanolamine **7**. As both ends of the molecule have nucleophilic sites, selective protection of one moiety would allow the other to be joined to the chloro-adenosine.

For the protection of the amine, the base stable protecting group, Boc-anhydride (Boc₂O), was used.¹⁹³ Although the hydroxyl is also potentially able to act as a nucleophile to the carbamate, the amine's much stronger nucleophilicity at the pH of the reaction would afford

the required chemoselectivity. To accelerate the reaction, a catalytic amount of the acyl carrier dimethylamino pyridine (DMAP) was added to the reaction.

^1H NMR spectroscopy confirmed that the carbamate **8** was synthesised by the appearance of an identifying upfield singlet, integrating for 9 protons, which corresponded to the *tert*-butyl group.

In contrast to the *N*-selective protection offered by Boc_2O , *tert*-butyldimethyl silyl chloride (TBDMSCl) was used to protect the hydroxyl.¹⁹⁴ The chemoselectivity of the silyl towards the oxygen is facilitated by the hard-hard interaction between the two atoms.

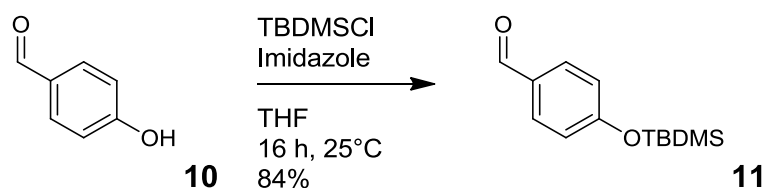
As hydrochloric acid is released in this reaction and would remove this acid labile protecting group, imidazole was added. During the work-up of the reaction, excess imidazole could not be removed using the conventional method of protonation as this would also have protonated the amine on the product. Instead, column chromatography was used for purification. After the isolation of the product **9**, ^1H NMR spectroscopy confirmed the successful synthesis of the silyl protected ethanolamine by the appearance of two upfield singlets, integrating for 6 and 9 protons.

Both the *N*- and *O*-protected groups were ready for attachment to the chloro-adenosine **4**.

Aldehyde to thiol conversion

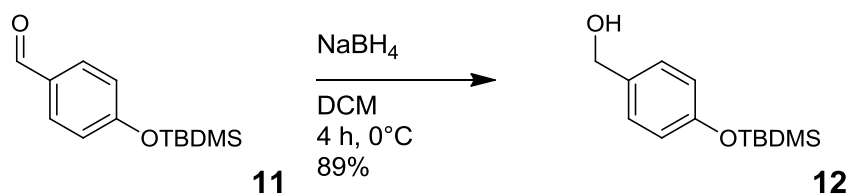
For the natural substrate of SpdSyn, dcAdoMet, the linking hetero atom between the adenosine and the aminopropyl moiety is a sulphur atom. In an effort to introduce this functionality, we developed a reaction sequence to convert an aldehyde into a thiol which could be joined to the chloro-adenosine **4**. This entailed the reduction of 4-hydroxybenzaldehyde **10** to the respective hydroxyl. Using a Mitsunobu reaction, a reaction which turns an acidic heteroatom into a leaving group, the hydroxyl was to be replaced with a thioester. Finally, the reduction of that thioester would then yield the respective thiol. The phenolic hydroxyl, however, had to be protected throughout the entire sequence.

Scheme 5-10 Synthesis of 11

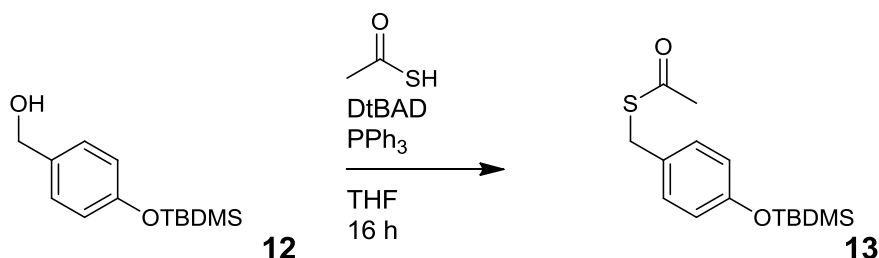


Firstly, using TBDMSCl, the protection of the alcohol **10** was successfully performed in good yield.¹⁹⁵

Analysis by ^1H NMR spectroscopy indicated the presence of the characteristic two upfield singlets that integrate for 6 and 9 protons.

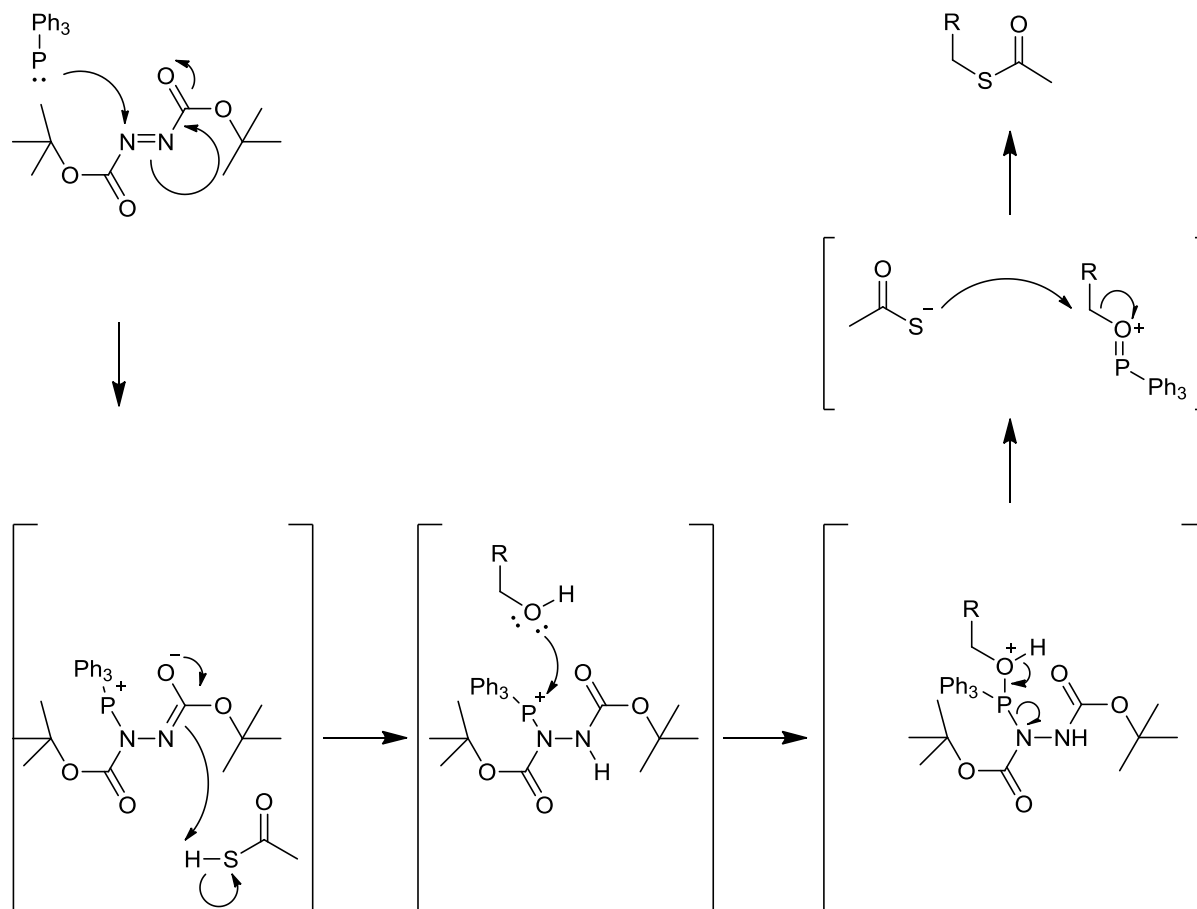
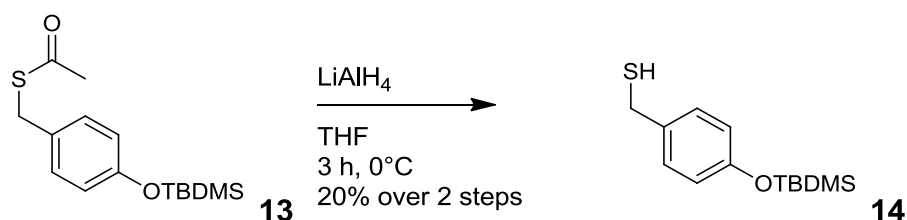
Scheme 5-11 Synthesis of 12

The reactive aldehyde **11** was then reduced using the mild reducing agent NaBH_4 .¹⁹⁶ The successful reduction to the alcohol **12** was confirmed by the disappearance of the downfield aldehyde singlet and the appearance of both a triplet corresponding to the hydroxylic proton and a doublet to correspond to the two benzylic protons.

Scheme 5-12 Synthesis of 13

To convert the alcohol into a leaving group, a Mitsunobu reaction was used in conjunction with thioacetic acid to afford a thioester.¹⁹⁷ The reaction mechanism is presented in Scheme 5-13. Although it wasn't relevant in this case, it is interesting to note that the step where the alcohol carbon is attacked by the nucleophilic thioate, proceeds via an $\text{S}_{\text{N}}2$ mechanism.

^1H NMR spectroscopy of the crude material confirmed that the thioester **13** was synthesised due to the disappearance of the upfield signal corresponding to the hydroxyl and the appearance of a singlet which integrated for 3 protons.

Scheme 5-13 Mitsunobu reaction mechanism¹⁹⁸Scheme 5-14 Synthesis of **14**

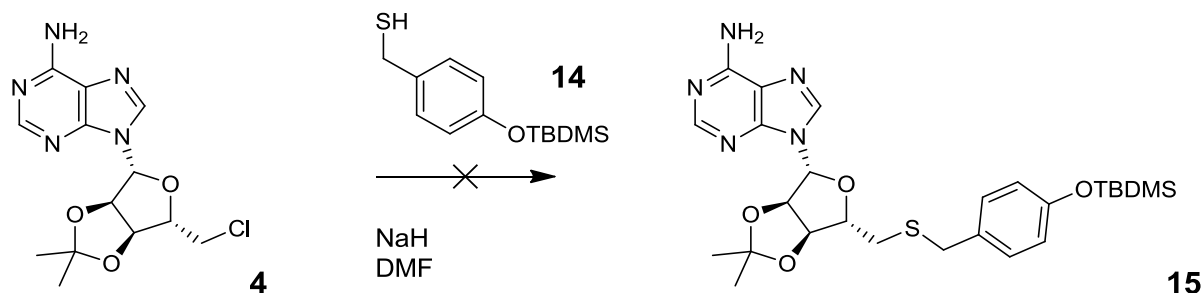
To reduce the thioester, a functionality that is far less susceptible to reduction than aldehydes or ketones, NaBH₄ was unsuitable. The much stronger reducing agent, lithium aluminum hydride, was chosen owing to its higher reactivity.¹⁹⁷ The reaction mechanism involves the formation of a complex with the carbonyl to facilitate hydride insertion. After an acidic work-up with a saturated solution of ammonium chloride, the thiol **14** was isolated in modest yield over two steps.

The product was characterised by ¹H NMR spectroscopy which observed that the methyl protons of the thioester have disappeared and that a triplet appeared, corresponding to the thiol, which integrated for one proton. Further, a doublet was observed which corresponded to the methylene protons with a coupling constant that matched the triplet corresponding to the thiol.

Synthesis – Linking attempts 1

After the completion of these five potential RHS molecules, the task at hand was to establish a protocol to link the molecules to the chloro-nucleoside **4**. To confer nucleophilicity the acidic heteroatoms, they would be deprotonated with a suitable strong base.

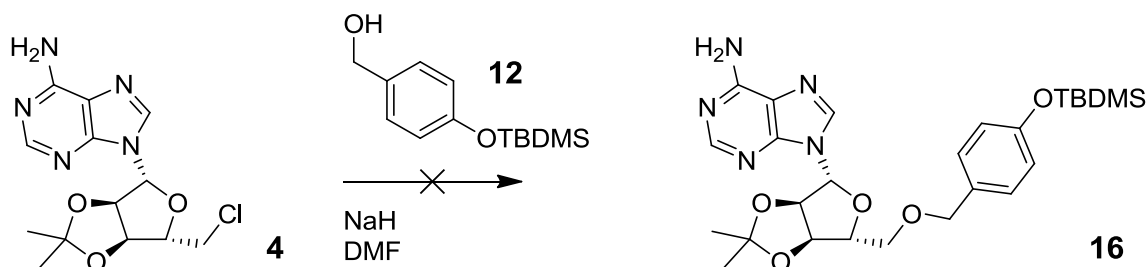
Scheme 5-15 Attempted synthesis of 15



The first target compound for the nucleophilic addition would be **15**. This was to be accomplished using a method analogous to Robins *et al* who synthesised a thioether using a thiol and chloroadenosine.¹⁹⁹

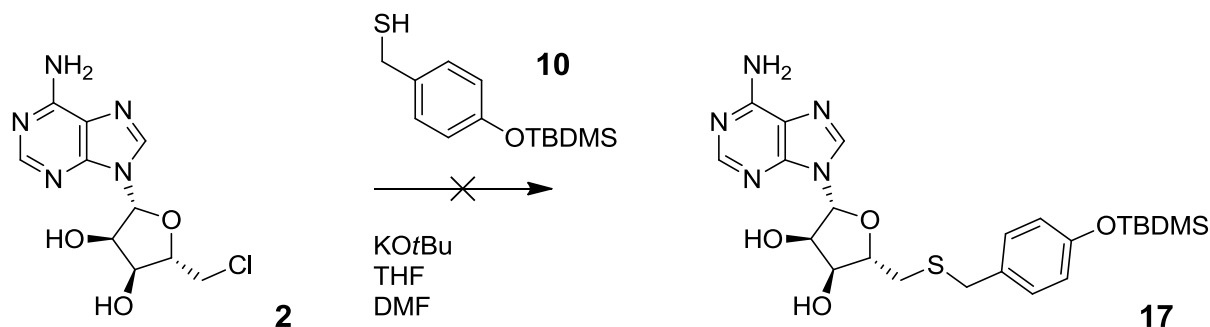
The thiol was added to an ice cold solution of DMF and treated with a 10% molar excess of NaH. After the anion was allowed to generate, the chloro-adenosine was introduced and the temperature increased. The reaction was monitored by TLC for an extended period, but no new UV-active compounds were evident and the reaction was abandoned.

Scheme 5-16 Attempted synthesis of 16



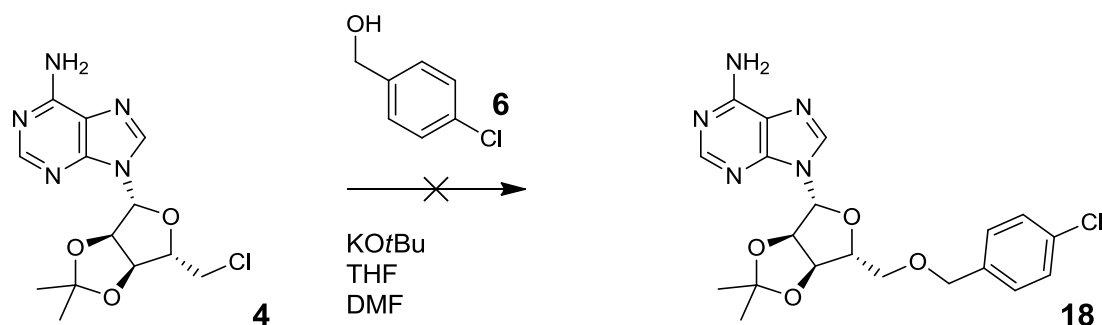
The reaction was setup again using similar conditions, this time using **12** to have oxygen as the nucleophile. Again, after monitoring the reaction for an extended time via TLC, no new products formed and the reaction was discarded.

Scheme 5-17 Attempted synthesis of 17



After the two failures, the NaH base was substituted for potassium *tert*-butoxide. To adhere to the conditions used by Kung *et al* as closely as possible,²⁰⁰ chloroadenosine without the ketal protection group was used. The reaction was monitored with TLC but once again, no products formed.

Scheme 5-18 Attempted synthesis of 18



In a final attempt to create an ether linkage, potassium *tert*-butoxide was added to the benzylic alcohol **06**. Unfortunately, in this last attempt, no possible products were detected for this reaction either. A possible explanation for these results could possibly include the presence of water in the reaction or inactive reagents. The latter explanation is unlikely as the addition of sodium hydride to the solution evolved hydrogen gas and the potassium *tert*-butoxide was freshly prepared by sublimation.

At this point, the attempts to synthesise thioether bonds were abandoned via the chloroadenosine. Instead, we would now reverse the strategy and introduce nucleophilicity to the adenosine **3**.

Conclusion

After optimisation, an electrophilic adenosine was successfully synthesised using thionyl chloride and pyridine in acetonitrile in moderate yield. Additionally, five functionalities were also prepared that were planned to be coupled to the chloro-adenosine. Numerous attempts were made to this aim, but were unfortunately unsuccessful for unknown reasons.

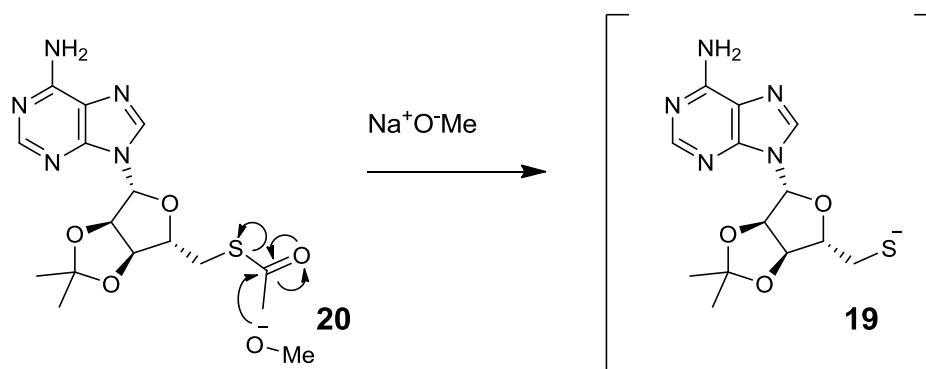
It would have been worthwhile to try alternative functionalities to introduce electrophilicity to the adenosine. For example, introduction of a mesylate or a triflate would have kept the primary alcohol oxygen in place without the harsh acidic conditions associated with the introduction of the chloride. Although there is a primary nitrogen present, its nucleophilicity is reduced compared to the hydroxyl oxygen, owing to its participation to the adjacent aromatic system. Furthermore, these alternative functionalities would also confer much stronger electrophilicity at the terminal carbon.¹⁹⁰

6. Spermidine Synthase – Nucleophilic adenosine

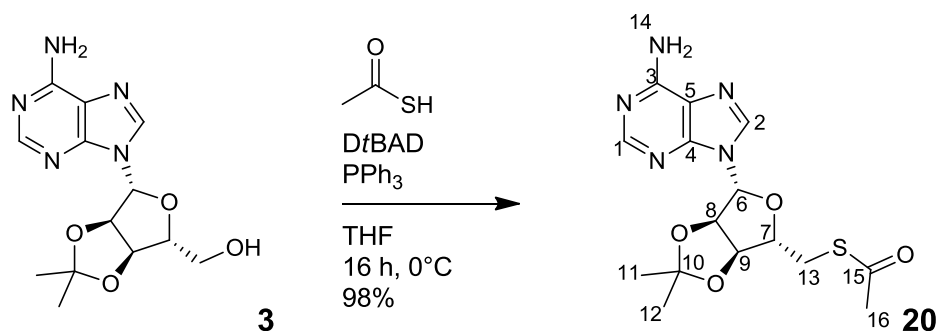
Synthesis – Thioester adenosine

Turning the strategy around, the left hand side of the molecule can also be turned into the nucleophile. Pignot *et al* successfully used adenosine derivatised with a thio-ester as a masked nucleophile to create thio-ether linkages to electrophilic molecules.²⁰¹ This nucleophilic thiolate anion **19** was prepared *in situ* by the addition of sodium methoxide to the thionucleoside **20** (refer to Scheme 6-1).

Scheme 6-1 Mechanism of nucleophilic conversion of thionucleoside via sodium methoxide²⁰¹



Scheme 6-2 Synthesis of the thionucleoside **20**²⁰¹



The thionucleoside was synthesised by a Mitsunobu reaction from the ketal-protected adenosine using the procedure from Pignot *et al.*²⁰¹ As an alteration for this reaction, instead of using the conventional Mitsunobu reagent, di-ethyl azodicarboxylate (DEAD), we used the di-*tert*-butyl azodicarboxylate (DtBAD) variant. This *tert*-butyl derivative is easily removed from the reaction by decomposing to volatile products in the presence of an acidic medium.¹⁹⁸ Using thioacetic acid, the corresponding thio-ester was synthesised by converting the primary hydroxy into a leaving group.

Mass spectrometry and NMR spectroscopy confirmed that the synthesis of the thioester was successful with a yield of 98%. All signals in the ^1H NMR spectrum were assigned, which also allowed all the signals in the ^{13}C NMR spectrum to be assigned, in conjunction with GHSQC (refer to Table 6-1). Once again the two protons at position 13, adjacent to the chiral centre, each presented unique signals.

Table 6-1 Assignment of 20 ^1H & ^{13}C NMR spectroscopy signals

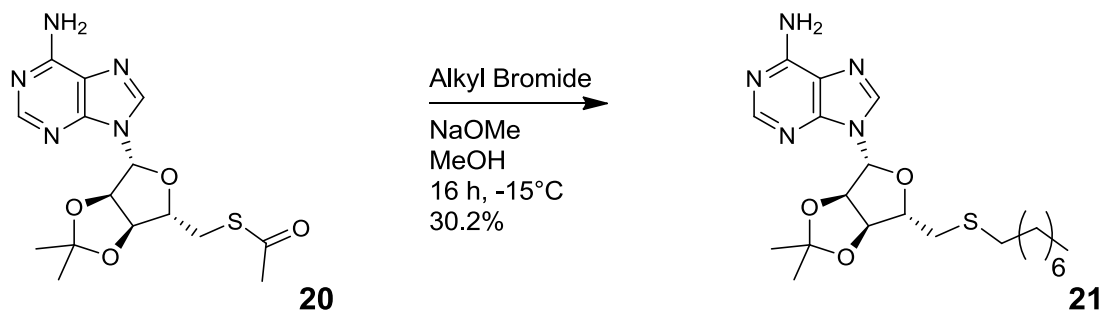
Atom position	^1H	^{13}C
15	–	194.5
3, 4 or 5	–	155.8
1 or 2	8.34 (s, 1H)	53.2
3, 4 or 5	–	149.2
1 or 2	7.89 (s, 1H)	139.9
3, 4 or 5	–	120.3
14	6.12 (s, 2H)	–
10	–	114.5
6	6.06 (d, $J = 2.1$ Hz, 1H)	90.9
8	5.50 (dd, $J = 6.4, 2.1$ Hz, 1H)	84.2
9	4.96 (dd, $J = 6.4, 3.1$ Hz, 1H)	83.7
7	4.33 (td, $J = 6.9, 3.1$ Hz, 1H)	86.1
13a	3.28 (dd, $J = 13.8, 7.2$ Hz, 1H)	31.3
13b	3.17 (dd, $J = 13.8, 6.7$ Hz, 1H)	
16	2.33 (s, 3H)	30.6
11 or 12	1.58 (s, 3H)	27.0
11 or 12	1.37 (s, 3H)	25.4

Synthesis – RHS moieties

Method development using alkyl chains

With the thionucleoside **20** in hand, we established the critical coupling reaction protocol as developed by Pignot *et al.*²⁰¹ To avoid complications, we used readily accessible monofunctionalised alkyl electrophiles.

Scheme 6-3 Synthesis of **21**



For our first reaction to join an alkyl group to the thionucleoside **20**, we used 1-bromo-octane. The simple alkyl chain was chosen due to its uncomplicated nature and was expected to be free from compounding side reactions

In our first attempt, we added the sodium methoxide to the stirred solution of the thionucleoside **20** at 0°C (refer to Table 6-2). After an hour of stirring, no change in the retention factor was observed by TLC. Nevertheless, we proceeded to add a large excess (10x) of the alkyl bromide, as the limiting reagent for this reaction was the thionucleoside. After leaving the reaction overnight, TLC analysis showed no evidence of the formation of a new compound. Fearing the possibility of an interfering elimination reaction, we halved the large excess of alkyl bromide to 5 equivalents for the subsequent attempt.

Table 6-2 Development of thioetherification procedure

Equivalents of alkyl bromide	Base (equivalents)	Solvent	Temperature	Yield
10	NaOMe (2.2)	MeOH	0°C	Unsuccessful
5	KOH (1.0)	DMF	0 to 55°C	Unsuccessful
5	NaOMe made <i>in situ</i> from Na metal (1.1)	MeOH	-15°C	Unsuccessful
5	NaOMe (1.0)	MeOH	-15°C	30.2%

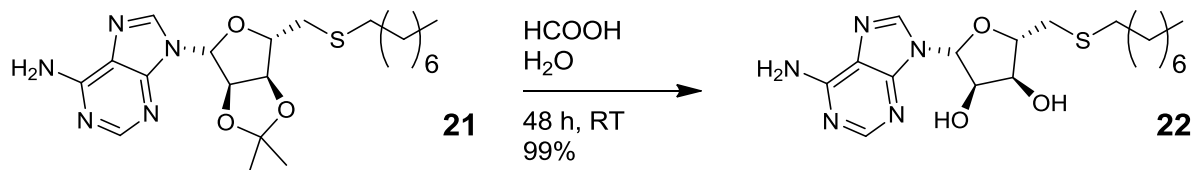
On this second attempt, we changed four parameters. Firstly, a smaller excess of the alkyl bromide was used and the solvent was exchanged for DMF, as it could help to stabilise the anion. Also, KOH was substituted as the base. To see whether it could help, the temperature

was also increased after sufficient time was allowed for the anion to generate. TLC analysis showed that this combination of parameters did not produce any new compounds either.

On the third attempt, we reverted back to the conditions in the original procedure by Pignot *et al.*²⁰¹ The reaction was started from -15°C in methanol and a stoichiometric amount of sodium metal was added to generate sodium methoxide *in situ*. Once again, no new compounds were observed by TLC.

Fearing that there could be water contamination in this particularly water sensitive reaction, all precautions possible were taken to remove all traces of water from the system for our next attempt. This included using solvent that was freshly distilled from CaH_2 .²⁰² Also, commercially produced NaOMe in methanol was also sourced for this reaction. After adding the methoxide base and allowing 15 minutes for the anion to generate, 5 equivalents of the alkyl bromide was added. The reaction was left to stir overnight and after TLC analysis, we discovered the presence of a new compound. After purification by column chromatography and lyophilisation, a yield of 30% was obtained for the compound **21**. ^1H , ^{13}C NMR spectroscopy and mass spectrometry then confirmed that the thioether bond was formed.

Scheme 6-4 Synthesis of **22**



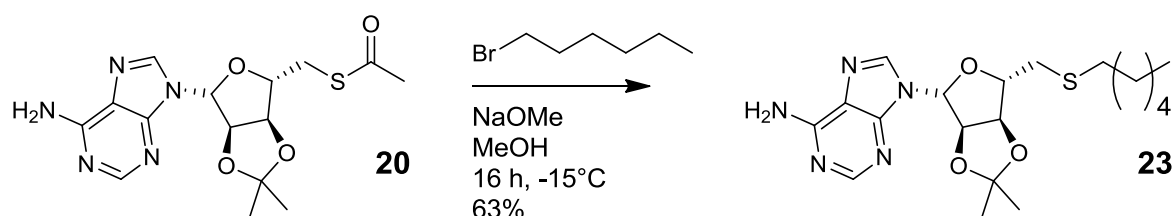
With the thioether **21** in hand, only the deprotection of the ketal was required to offer the first final product in this series. This was easily achieved through the use of formic acid and purified water.

After monitoring the reaction using TLC, the starting material was consumed after two days. As planned, all by-products and reagents of this reaction are volatile: formic acid, water and acetone. These were removed by a combination of rotary evaporation and high vacuum. ^1H NMR spectroscopy then confirmed that the two methyl peaks correlating to the ketal methyl protons disappeared. Further, ^{13}C NMR spectroscopy and mass-spectrometry confirmed that the isolated product was pure and agreed with expected results. After the deprotection of the hydroxides, the highly polar nature of this compound dictated that deuterated methanol was required for the NMR experiment as solvent. Due to this, the hydroxyls and the amines were subject to deuterium exchange and no proton signals for these moieties were observed. Also, owing to the extreme hygroscopic nature of methanol, water contamination of the solvent is unavoidable. To circumvent this, a signal is introduced into the NMR pulse sequence to saturate the water signal. Performing this water

Chapter 6 – Spermidine Synthase (Nucleophilic adenosine)

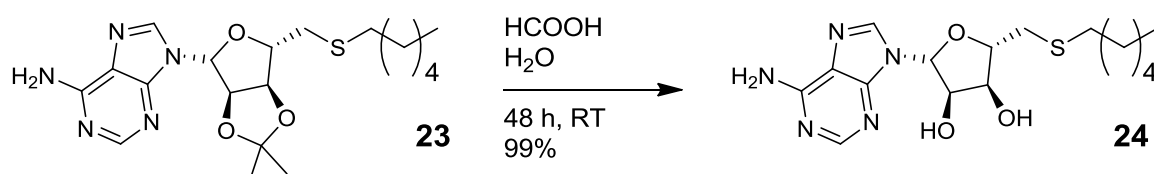
suppression drastically reduced the size of the peaks in representing water and was applied whenever deuterated methanol was used.

We had our first final compound in hand with a yield of 11.75 mg. Armed with the sound confidence that our method works, we set to work to join the remaining three alkyl bromides to the thionucleoside.

Scheme 6-5 Synthesis of 23

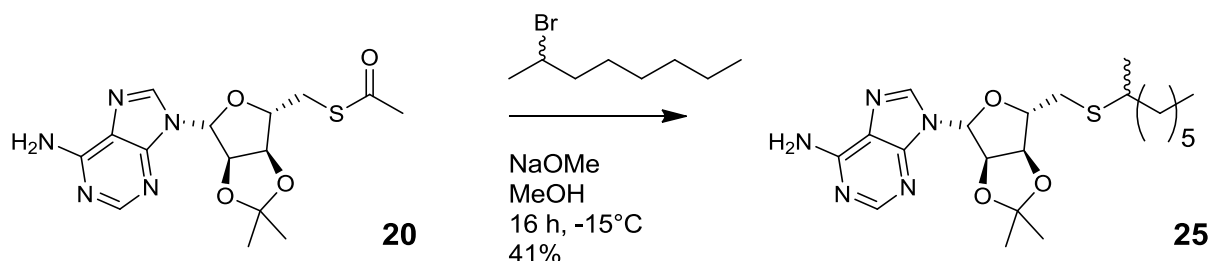
Similar conditions for the synthesis of octyl **21** were used to prepare the hexyl derivative **23**. Again using strict anaerobic conditions, the alkylation of the thionucleoside was performed.

After quenching the reaction with water, purification using chromatography and lyophilisation afforded the pure product. ¹H and ¹³C NMR spectroscopy and mass-spectrometry confirmed that the hexane derivatised thionucleoside **23** was formed.

Scheme 6-6 Synthesis of 24

After the acid catalysed deprotection of the vicinal hydroxides, the second pure product **24** in the series created.

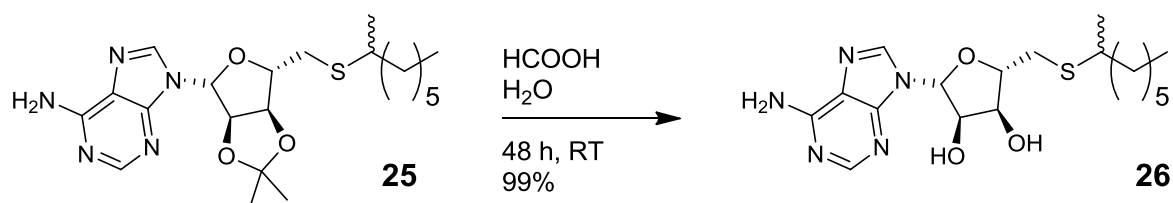
¹H NMR and ¹³C NMR spectroscopy in conjunction with mass spectrometry confirmed the successful synthesis of the second compound in the series.

Scheme 6-7 Synthesis of 25

We then proceeded to derivatise secondary halogenated alkanes. Although the secondary halogenated carbon would be slightly less electrophilic, the reaction went without difficulty.

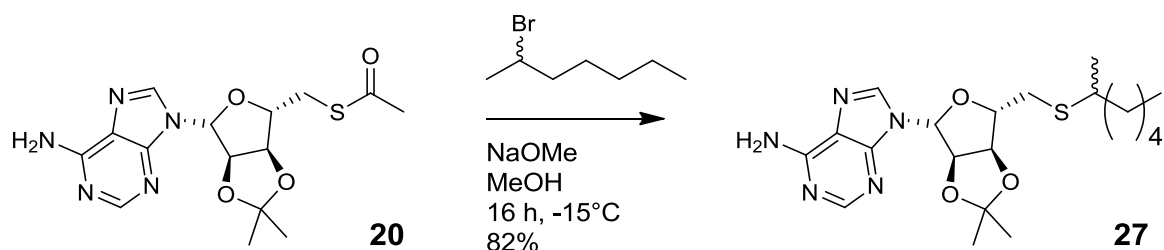
Chapter 6 – Spermidine Synthase (Nucleophilic adenosine)

Mass-spectrometry and NMR spectroscopy confirmed that that synthesis proceeded as planned. However, a racemic mixture of 2-bromo-octanol was used and this was incorporated into the thioether **20** to create diastereomers. Therefore, atoms that are in close proximity to the racemic chiral centre are in slightly different chemical environments as the corresponding atoms in the opposite diastereomer and these differences are observed by NMR spectroscopy. These differences are most evident on the atoms adjacent to the chiral centre and become gradually less pronounced farther away. For this reason, ^1H NMR signal multiplicities could not be resolved in all instances and assignments could not be performed by deduction from the coupling constants. Therefore, most signals were reported as multiplets and assignments were based on the knowledge of the previously obtained spectra. Additionally, ^{13}C NMR spectra presented similar complications. As a result, spectral signals could not be unambiguously assigned to the structure. This complication persisted for the remaining diastereomeric compounds in this sequence. As these diastereomers did not separate using conventional chromatographical techniques, further separation attempts would only be attempted should there be significant biological activity.

Scheme 6-8 Synthesis of 26

From here, the ketal was removed using formic acid and water to afford the two free hydroxyls.

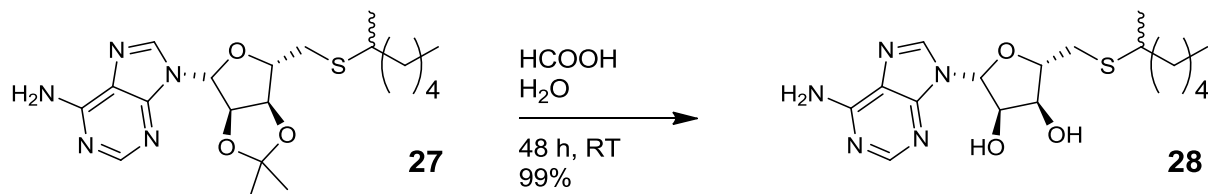
^1H and ^{13}C NMR spectroscopy and mass-spectrometry confirmed that the third compound, **26**, in the series was successfully formed. This pair of diastereomers was submitted for testing.

Scheme 6-9 Synthesis of 27

Our final alkylation was performed with 2-bromo-heptane to yield **27**.

Again, due to the incorporation of the racemate, two diastereomers of the thioether were created and were visible in both the ^1H and ^{13}C NMR spectra. Mass-spectrometry was also in agreement with the expected products.

Scheme 6-10 Synthesis of 28



Once again, the final deprotection of the ketal **27** was performed over 2 days using formic acid and water.

After purification, ^1H and ^{13}C NMR spectroscopy and mass-spectrometry confirmed the successful deprotection of the ketal and the successful formation of the final alkylation product **28**.

AdoDATO and 4MCHA hybrid

Armed with a working method to derivatise the thionucleoside and with several inhibitors in hand, we would now design and synthesise inhibitors with additional functionality to interact with the binding pocket. Looking to the two known inhibitors for inspiration, we attempted to use characteristics of both molecules in a single hybrid molecule.

To do this, we superimposed the two known inhibitors into the same co-ordinate space (refer to Figure 6-1). We observed that the far amine of AdoDATO and the amine from 4MCHA both interact with the protein in same region, which we shall refer to as the distal binding pocket. We would also refer to the interaction zone of the second amine of AdoDATO as the proximal binding pocket. Because there was a region of overlap between the two known inhibitors, we would combine them into our first iteration of a hybrid molecule. This resultant molecule **29** contained synthetically challenging chiral centres but fitted very well within the binding site with a CDOCKER score of 47, compared to the potent inhibitor AdoDATO with a score of 82. 4MCHA could not be used as reference due to the difference of the inhibitor's size and interaction regions.

Chapter 6 – Spermidine Synthase (Nucleophilic adenosine)

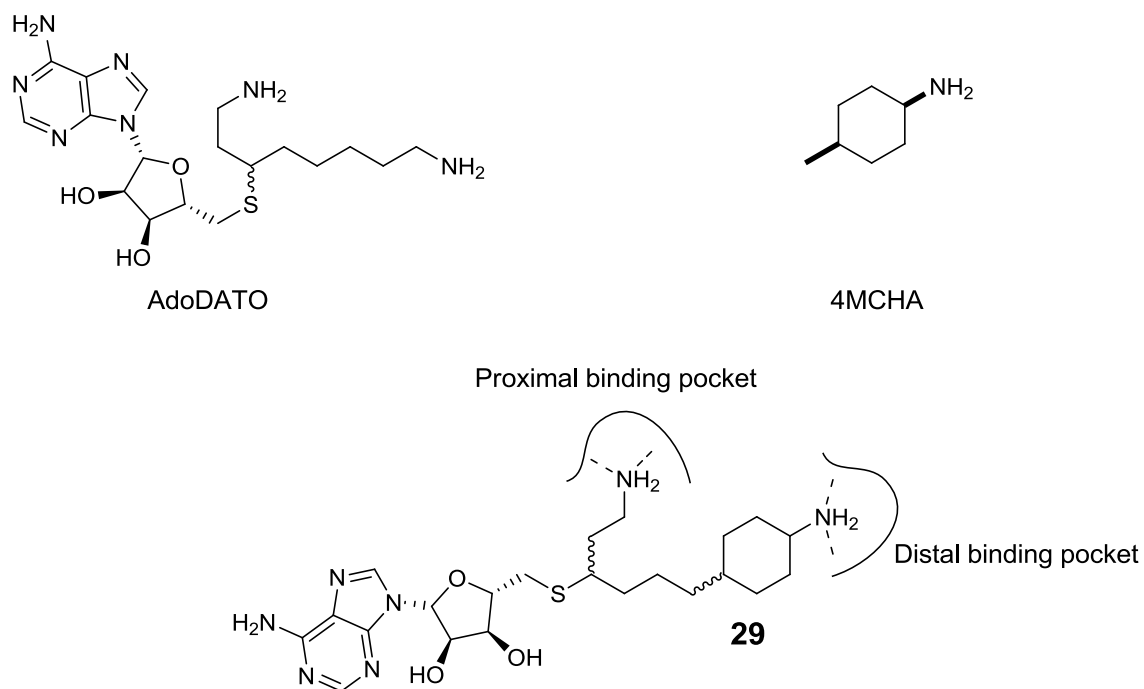


Figure 6-1 First iteration of a hybrid molecule

To negate the one chiral centre, the cyclohexane ring was aromatised to an aniline derivative **30** (Figure 6-2). After docking this structure into the protein, it was evident that there could be a pi-stacking interaction with a tyrosine residue located directly above the aniline (Figure 6-3). This improved the CDOCKER score to 65 for aniline derivative **30**.

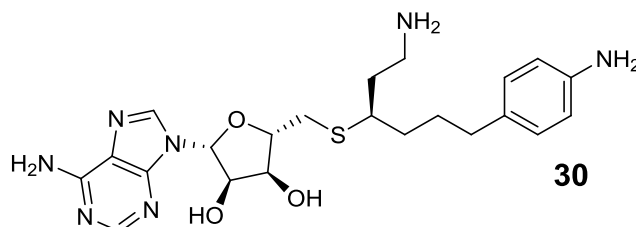


Figure 6-2 Ideal hybrid 30

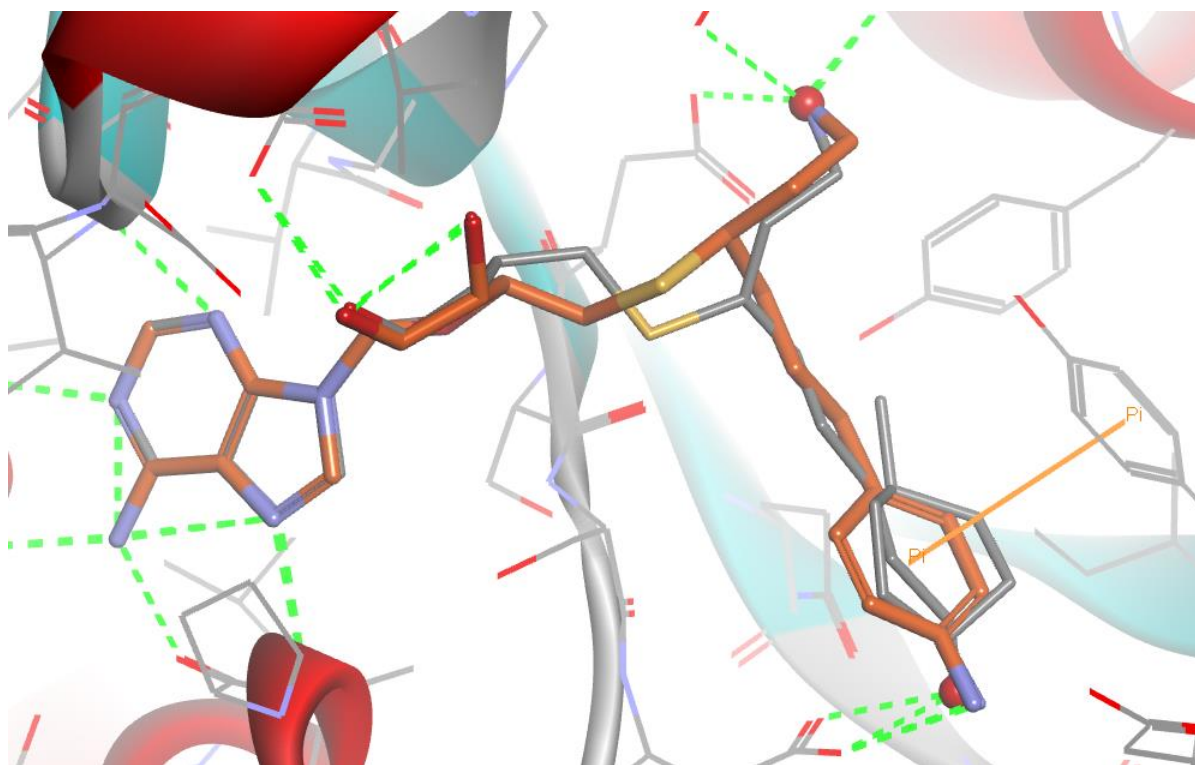
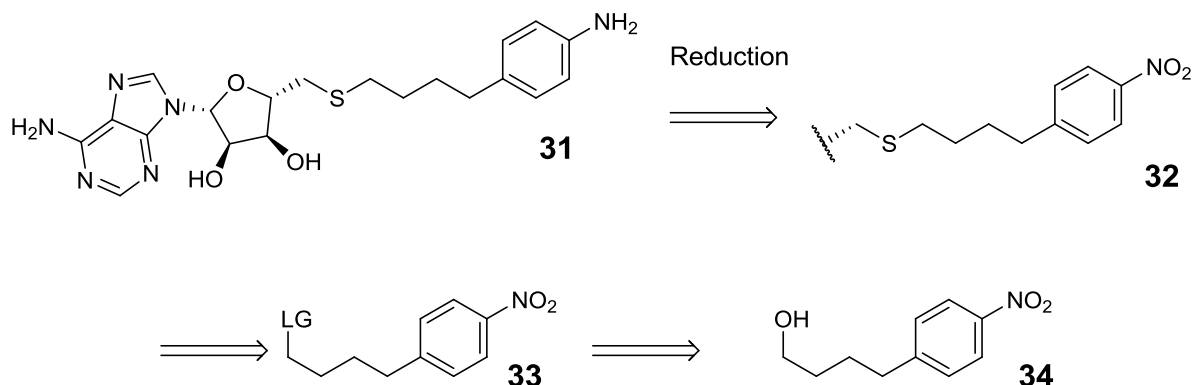


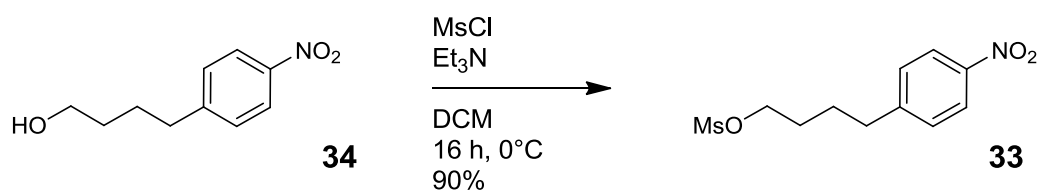
Figure 6-3 Proposed hybrid of AdoDATO and 4MCHA docked into the active site of SpdSyn. The two strong inhibitors, AdoDATO and 4MCHA, are presented here in the active site of SpdSyn. The aniline derivative **30** (carbons in orange) incorporates interactions of both of the known inhibitors.

Even though this molecule looked like a promising inhibitor, the incorporation of both segments which interact with the two binding pockets was synthetically unfeasible at the time. To quicken our development, the propylamine segment was removed from the design which yielded a molecule, **31**, which would be easily accessible, yet still contain the main interaction criteria (see Scheme 6-11). This molecule would be made by joining the nucleophilic thionucleoside to an aniline derivative with a suitable electrophilic centre. Because the amine could potentially cause unwanted effects during the linking reaction, it would be masked as a nitro group before being reduced to the amine (using the Pd/C) as shown in the disconnection to **32**.

Scheme 6-11 Disconnection of 30



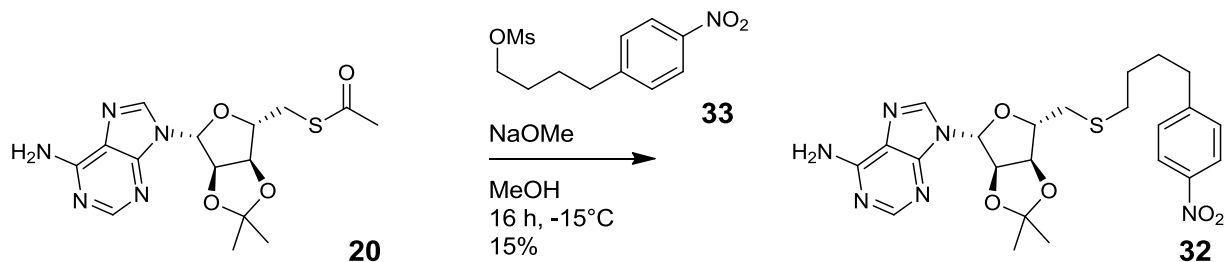
Scheme 6-12 Synthesis of 34



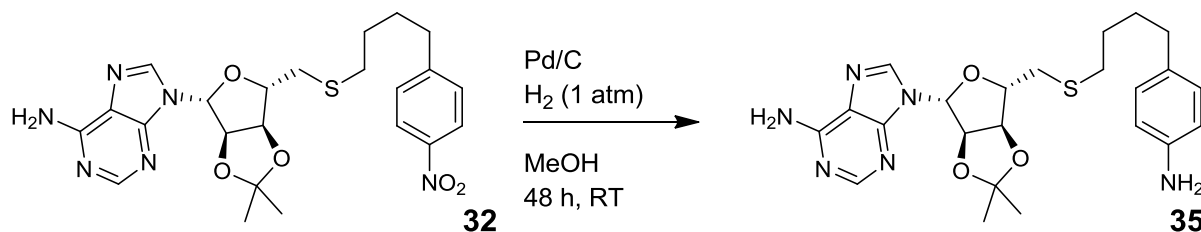
An efficient way to turn a hydroxyl into a leaving group is by its conversion to a mesylate and this reaction was used for the derivatisation of 4-(4-nitrophenyl)butan-1-ol **34** in excellent yield.²⁰³

After purification, ¹H NMR spectroscopy indicated the presence of an upfield singlet which integrated for 3 protons, which confirmed that the mesylate was obtained in good yield.

Scheme 6-13 Synthesis of 32



The mesylate was then ready for coupling to the thionucleoside using the previously established method (refer to Scheme 6-5). Although only a poor yield was obtained, ¹H and ¹³C NMR spectroscopy, and mass spectrometry confirmed that the nitrobenzene nucleoside was obtained.

Scheme 6-14 Synthesis of 35

To reduce the nitro group to the respective amine, palladium on carbon and hydrogen gas was used. After 3 days, a new product was observed by TLC analysis, which also tested positive for ninhydrin, which suggested that the amine had formed.

Unfortunately, the completion of this promising compound **35** was never realised due to time constraints.

Tyrosine derivative

In an alternative attempt to produce a hybrid with a phenyl moiety, we would incorporate tyrosine into the thionucleoside to yield the tyrosine derivative **36** (Figure 6-4).

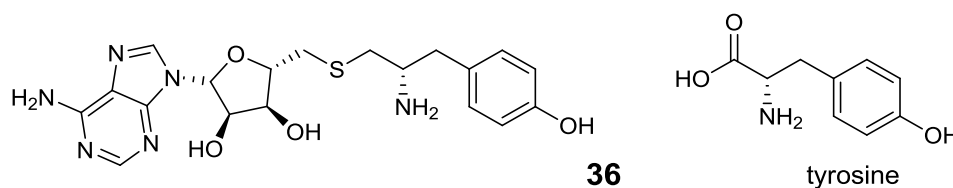


Figure 6-4 Proposed tyrosine hybrid derivative 36 and tyrosine

Although tyrosine has a hydroxyl at the *para*- position (as opposed to an amine), the hydroxyl can also act as a hydrogen bond donor if rotated to the correct position. Modelling suggested the possibility of a new hydrogen bond interaction that could be made to the protein via the tyrosine amine (Figure 6-5). This tyrosine derivative **36** obtained a CDocker score of 49 (for ease of reference AdoDATO has a CDocker score of 83).

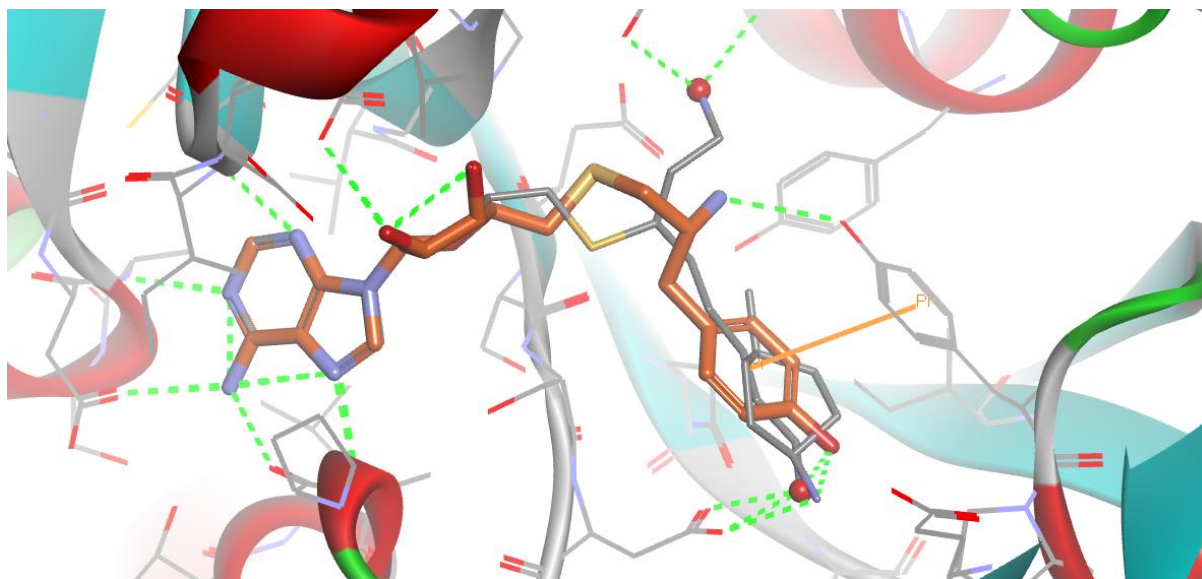
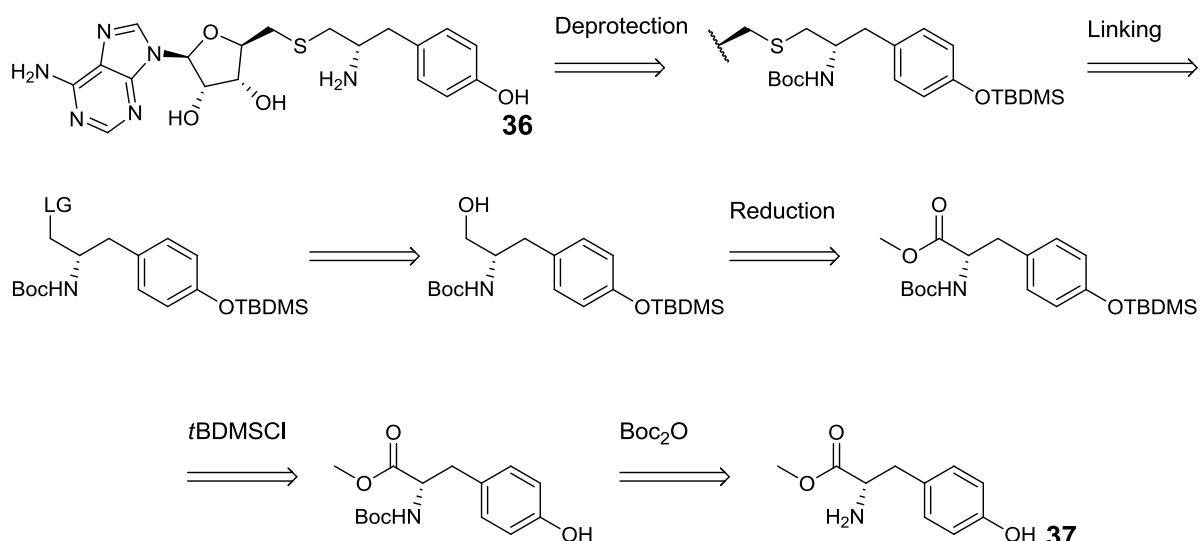
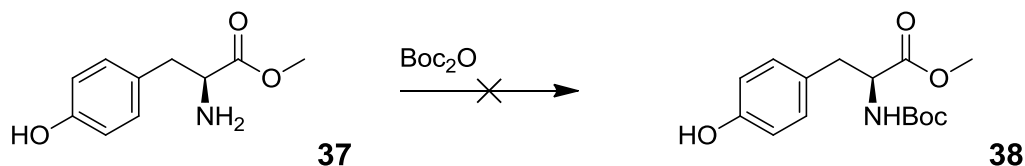


Figure 6-5 Proposed hybrid of 4MCHA, **36**, docked into the active site of SpdSyn. Hydrogen bonding interactions are shown in green, whilst the orange line indicates a possible pi-interaction. For reference, AdoDATO is also shown (grey carbons).

To incorporate tyrosine into the thionucleoside, a disconnection was planned that would use esterified tyrosine **37** as the starting material (see Scheme 6-15). Both the amine and phenolic oxygen would be protected with the base stable protecting groups, Boc and *t*BDMS for the amine and hydroxyl, respectively. An appropriate leaving group for the thiolation can then be installed on the ester carbon by reducing it to the hydroxyl using LAH and which will then be mesylated. Ultimately, the protecting groups would have to be removed.

Scheme 6-15 Disconnection of **36**

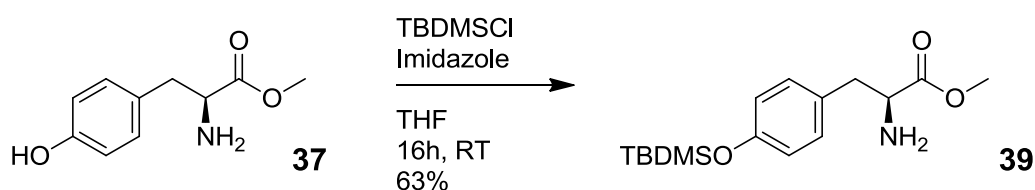


Scheme 6-16 Synthesis of 37

The synthesis commenced with the Boc protection of the amine using Boc_2O in acetonitrile in combination with the acyl transfer catalyst DMAP.²⁰⁴ Unfortunately, no new analytes were observed using TLC. After the misfortune, a literature search found that this reaction may also be performed in THF to excellent yield using the same catalyst.²⁰⁵ Again, the expected product was not formed. Substituting the DMAP for TEA, another group synthesised the carbamate using DCM as solvent.²⁰⁶ Unfortunately, the final attempt in ethanol using NaHCO_3 also did not work.²⁰⁷ For a tabulation of these reaction conditions attempted, refer to Table 6-3.

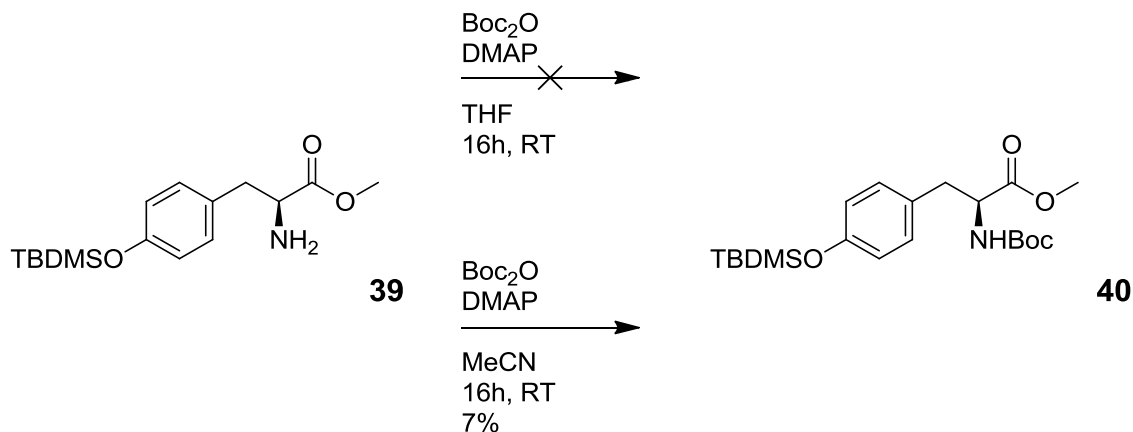
Table 6-3 Conditions attempted for the synthesis of 38

Catalyst / Base	Solvent
DMAP	MeCN
DMAP	THF
TEA	DCM
NaHCO_3	EtOH

Scheme 6-17 Synthesis of 39

In an effort to circumvent the problematic reaction, the protection of the hydroxyl was first attempted by means of *tert*-butyl-dimethyl-silyl chloride (TBDMSCl).²⁰⁸

The synthesis of the silyl ether was confirmed with the appearance of two singlet signals which integrated for 9 and 6 protons in the ^1H NMR spectrum.

Scheme 6-18 Synthesis of **40**

This allowed us to re-attempt the amine protection in a slightly altered environment. The reaction was performed in THF, which resulted in a number of side reactions, the products of which were not isolated. The solvent was then changed to acetonitrile, which reduced the by-products to only 3. After purification, one of these products was confirmed by ^1H NMR spectroscopy to be the desired carbamate with the appearance of an upfield singlet which integrated for 9 protons. Unfortunately, only a trace amount of desired product **40** was obtained from this reaction. At this point, this synthetic route was abandoned for more promising compounds.

Triazole derivatives

Design

Another molecule was designed to incorporate an aromatic ring, this time a triazole. Triazoles are both synthetically easily accessible and have good drug-likeness. The triazole moiety would be well accommodated in the same region as 4MCHA, also possibly with a pi-interaction with the above-lying tyrosine residue. The triazole would be substituted with an amino-methyl group to be able to interact with the distal cation binding pocket (see Figure 6-6). The triazole derivative **41** that was planned for synthesis was docked into the prepared protein structure with a CDOCKER score of 43 (for ease of reference AdoDATO has a CDOCKER score of 83).

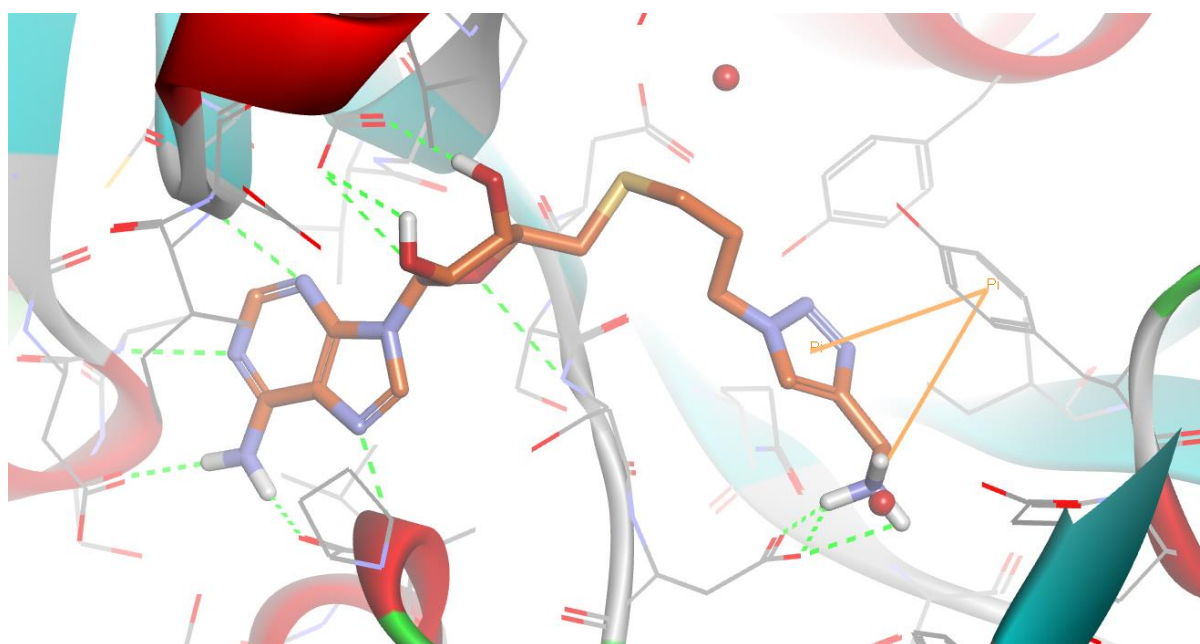


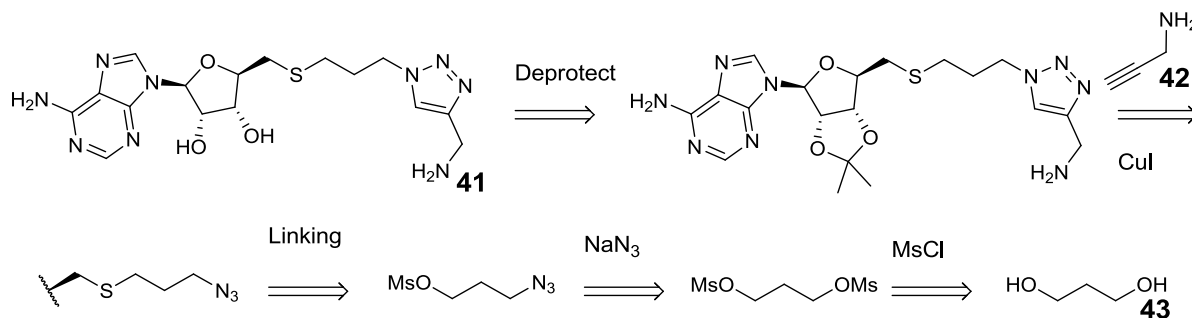
Figure 6-6 The 3 carbon linked triazole derivative, docked into the active site. The red dots designate the cation binding sites.

Planning

The disconnection was planned as follows (Scheme 6-19). The triazole could be formed using a Huisgen 1,3-dipolar cycloaddition from an azide. Using copper as catalyst would ensure that only the desired 1,4 regioisomer **41** would be formed with propargylamine **42** after deprotection.²⁰⁹ The azide would be linked to the nucleophilic adenosine via an electrophilic centre, easily accessed from a hydroxyl. The azide functionality would be accessed by the reaction of sodium azide and a hydroxyl. This concluded the disconnection sequence with propane-1,3-diol **43** as starting material.

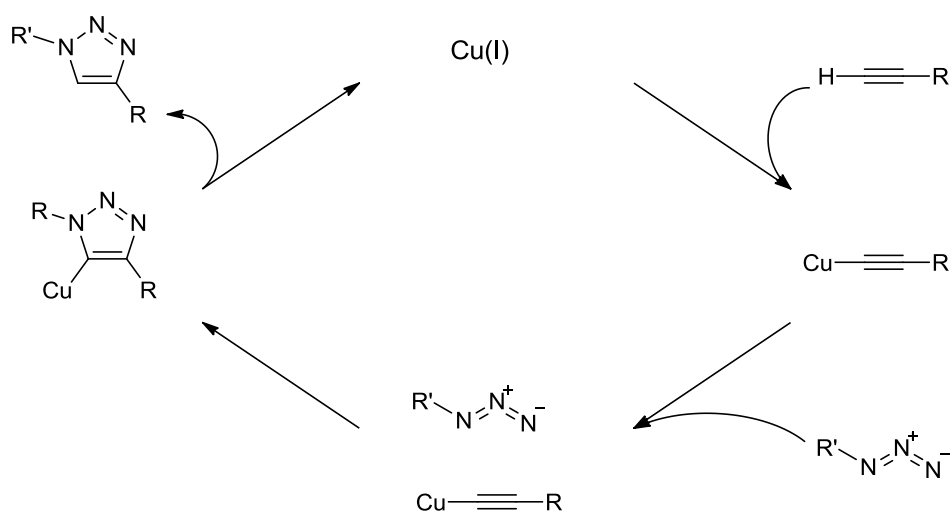
Chapter 6 – Spermidine Synthase (Nucleophilic adenosine)

Scheme 6-19 Synthesis pathway to the proposed triazole derivative 41



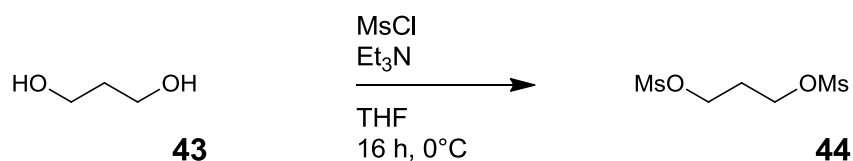
Scheme 6-20 Mechanism of the copper catalysed Huisgen 1,3-dipolar cycloaddition

The exact mechanism of the reaction is still unknown. However, it is known that copper is able to co-ordinate to the alkyne after the acidic proton has been abstracted with a suitable base.²¹⁰



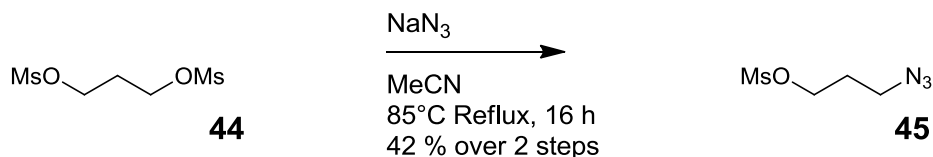
Synthesis

Scheme 6-21 Synthesis of 44

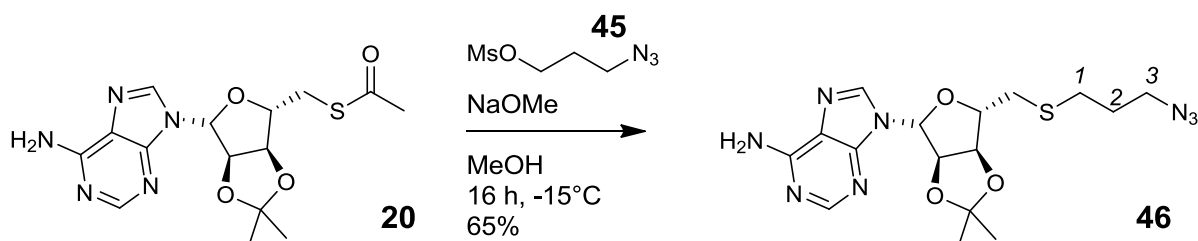


The construction of the three carbon linked triazole was started with the *bis*-mesylation of 1,3-dipropyl azide by the use of mesylchloride.²¹¹

After the reaction was left to stir overnight, the salt was removed by filtration and the volatiles were removed using flash distillation. The crude product was then used to create the azide without further purification.

Scheme 6-22 Synthesis of 45

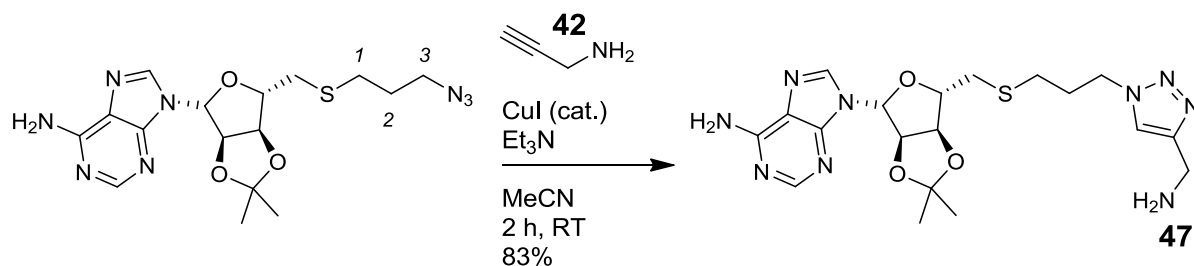
To make the azide, a stoichiometric amount of NaN_3 was added to the crude dimesylate and heated at reflux overnight. After the reaction was completed, TLC was used to examine the products formed. To visualise the azide, we used two stains. One, reliant on a mini Click-chemistry reaction, proved to be very ineffective.²¹² Another, reliant on a mini Staudinger reaction, proved to be very sensitive.²¹³ The latter entailed dipping the TLC plate in a PPh_3 solution to reduce the azide to the amine and then dipping the plate in ninhydrin to visualise the amine. After column chromatography, the yield was 42% over the two steps. We were now ready to join this molecule to the thionucleoside. The ^1H NMR spectrum was used to confirm that the molecule was expected by the disappearance of one mesyl group and change in chemical shifts of the signals. Additional characterisation was not performed as the compound is not novel.²¹⁴

Scheme 6-23 Synthesis of 46

With our established method (refer to Scheme 6-5), a slight excess of NaOMe was used to generate the anion **19** at -15°C for 1 hour. After this, the azide **45** was added to stir overnight.

The reaction was quenched with a small amount of water. After purification using column chromatography, the compound was analysed by ^1H NMR, ^{13}C NMR and mass spectrometry. Two triplet signals, both of which integrated for two protons, were present in the ^1H NMR spectrum which correlated to the two atom positions indicated (1 and 3). Due to the increased electronegativity of nitrogen over sulphur, the more downfield signal was assigned to the protons of the carbon neighbouring the azide (3).

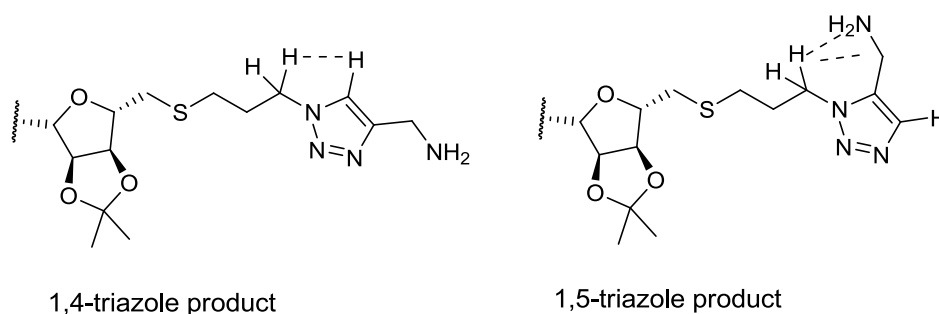
Scheme 6-24 Synthesis of 47



With the thioadenosine **20** suitably derivatised with the azide, we were ready for the copper catalysed Click reaction. To ensure that the 1,4 regioisomer would form, CuI was used as catalyst for the reaction. A slight excess of propargylamine **42** was added to ensure that the reaction went to completion.²¹⁵

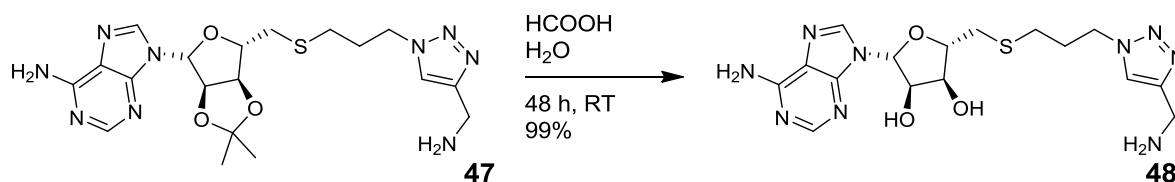
The reaction was monitored by TLC. After an hour, starting material was still evident in the reaction mixture, so additional propargyl amine was added. After another hour, no azide was observed. The reaction was then quenched with a saturated solution of NH_4Cl and purified using column chromatography to give a good yield. ^1H NMR spectroscopy analysis presented the opportunity for several observations:

1. The presence of an additional singlet signal in the aromatic region was observed. This was the primary indication of the successful formation of the triazole.
2. The signal that was assigned to the protons on the carbon adjacent to the sulphur for the preceding reaction (1), had very small difference in the chemical shift (2.5 vs 2.4 ppm). In contrast, the protons present on the carbon adjacent to the azide (3), shifted much more as these were now adjacent to the much more strongly electron withdrawing triazole (3.3 vs 4.3 ppm). This observation provided additional confidence that the preceding assignment was correct.
3. Using copper to catalyse the Huisgen 1,3-dipolar cycloadditions, only the 1,4-disubstituted triazole was expected to form.²⁰⁹ In an effort to confirm that the 1,4-regioisomer had indeed formed, a NOESY NMR experiment was performed even though we were not too hopeful of conclusive results. The experiment was based on the principle that if the 1,4-isomer had formed as expected, that the protons adjacent to the triazole could be in close enough proximity to the aromatic proton at the 5' position to allow cross-coupling to be observed (drawn explicitly, see Figure 6-7). Unfortunately, no definitive “through space coupling” was observed for either of the two possible regioisomers.

**Figure 6-7** Two possible triazole regioisomers

If the 1,4-regioisomer had formed, through space coupling could have been observed between the triazole neighbouring protons and the aromatic proton. Alternatively, through space coupling could have been observed to either the amine or neighbouring carbons protons.

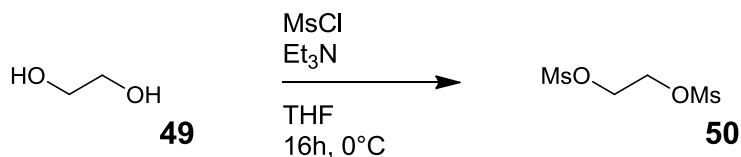
All 19 of the carbon atoms were accounted for in the ^{13}C NMR spectroscopy experiment. LCMS confirmed that the mass was as expected and UV absorbance indicated that the product was 97.8% pure. The final deprotection step could now be performed to remove the ketal.

Scheme 6-25 Synthesis of **48**

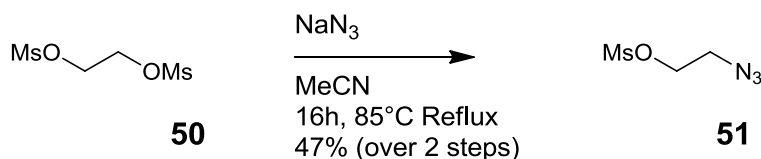
The triazole **47** was dissolved in formic acid and purified water to remove the ketal protecting group. After two days, no starting material could be detected by TLC. All the volatiles were then removed by a combination of vacuum distillation and high vacuum to yield the di-hydroxy nucleoside **48**. LCMS indicated that the resultant product was 89.7% pure by UV, which also confirmed that the mass was as expected.

^1H NMR spectroscopy confirmed the disappearance of the two singlets in the upfield region, which correlated to the methyl peaks of the ketal. ^{13}C NMR spectroscopy indicated the presence of 16 carbons, which also correlated with the structure. This compound was then submitted for enzymatic assay testing.

In an analogous sequence, these steps were repeated on a 2 carbon diol **49** to explore the number of linker atoms to the thiazole and amine.

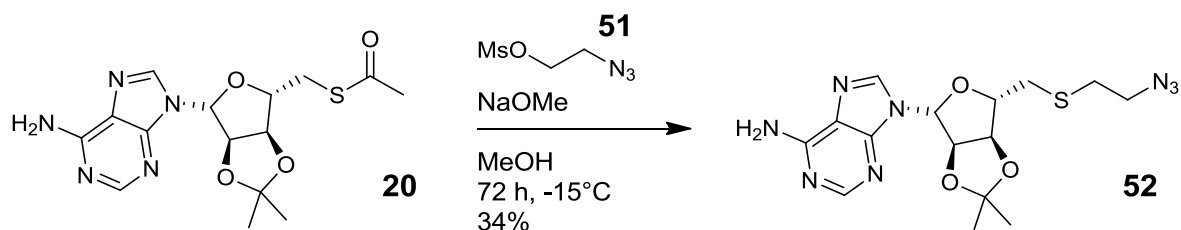
Scheme 6-26 Synthesis of 50

Once again, the sequence was started by the *bis*-mesylation of ethane-1,2-diol **49**.²¹¹ After the reaction was left to proceed overnight, the triethylamine salts were removed by filtration and the volatiles were removed by vacuum distillation. The compound was used in the following azidification reaction without purification.

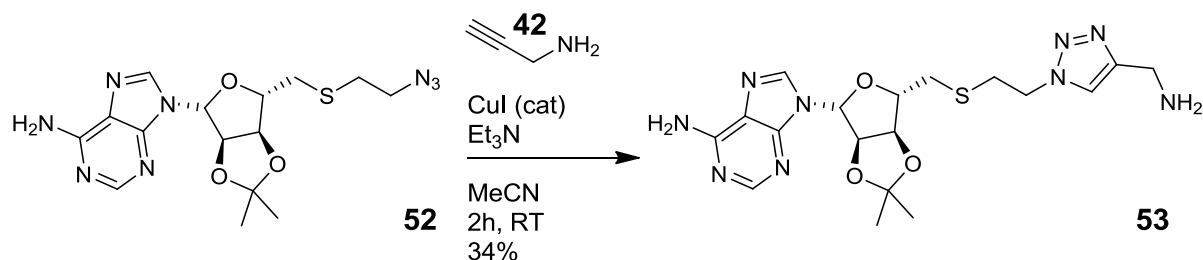
Scheme 6-27 Synthesis of 51

The azidification was performed under reflux conditions overnight.²¹¹ Because the di-azide could potentially be explosive, exactly one stoichiometric equivalent of sodium azide was added to the reaction, based on the mass of the diol.

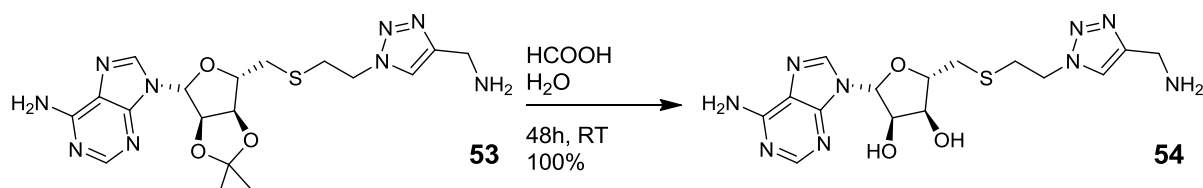
After purification using column chromatography, ¹H NMR spectroscopy confirmed that one mesyl group was present with the presence of an upfield singlet signal which integrated for 3 protons with no hydroxyl peak evident.

Scheme 6-28 Synthesis of 52

Again, a modest yield was obtained over the two steps to use in the reaction to couple the azide **51** to the thionucleoside **20**.²⁰¹ The reaction was performed in freshly distilled methanol, in which the nucleophile was prepared using sodium methoxide. After 3 days, the reaction was deemed to be complete, as all of the thionucleoside was consumed. After the product was isolated and purified, the obtained ¹H and ¹³C NMR spectra and the mass spectrum was consistent with the structure **52**.

Scheme 6-29 Synthesis of 53

The azide **52** would again be converted to a 1,4-triazole using Click chemistry. The reagents were added together with a final addition of a catalytic amount of copper iodide. The reaction was monitored by TLC and an additional portion of propargyl amine was added after an hour to facilitate the reaction. ^{13}C NMR spectroscopy data was consistent with the number of carbons and ^1H NMR spectroscopy indicated the appearance of a singlet in the aromatic region which would indicate the successful formation the triazole and the disappearance of the singlet in the upfield which would indicate the loss of the thionucleoside methyl group. All other signals in the ^1H NMR spectrum were assigned to the expected molecule **53**.

Scheme 6-30 Synthesis of 54

Only the ketal deprotection remained using a combination of formic acid and water. The reaction was left to stir for 2 days and monitored by TLC. When no more ketal **53** could be detected, the product was collected by removing all the volatiles. ^1H NMR spectroscopy confirmed the disappearance of two singlets in the upfield region, which would correspond to the ketal protons. Signals which would correspond to the hydroxyls would not be seen due to proton exchange with the deuterated methanol solvent used for the experiment. ^{13}C NMR and mass spectrometry data were also concordant with the structure and confirmed that the second triazole compound was complete.

Results

After the completion of the syntheses, the compounds were sent for enzymatic analysis.²¹⁶ The results of the tests are presented in Figure 6-8. The 4 alkyl derivatives, which are intended to occupy the hydrophobic region of the active site presented no inhibition. This can either be ascribed to experimental error or an unknown, possibly agonistic, mechanism.

Of the 5 compounds, the triazole derivative **48** presented the strongest inhibition activity, albeit marginal.

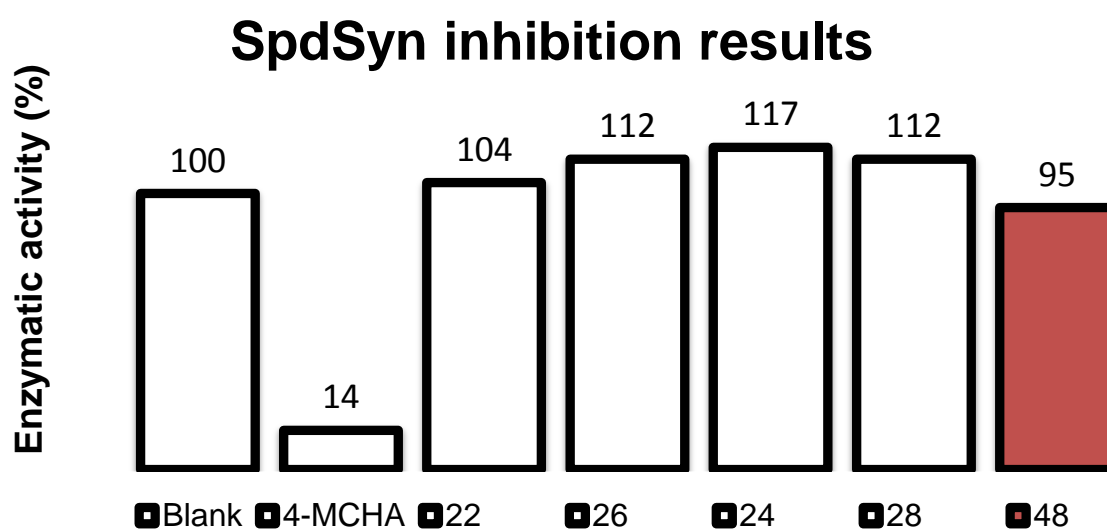


Figure 6-8 SpdSyn inhibition results

The enzymatic activity was conducted at 100 μM for each compound. The triazole derivative is highlighted in red. For reference, the potent inhibitor 4-MCHA is included.

Conclusion

Nucleophilicity was successfully conferred to adenosine by the incorporation of a thioester onto the molecule using a high yield Mitsunobu reaction.

By releasing the masked thiolate using sodium methoxide, 4 alkyl chains were also introduced to the adenosine. This set the stage to incorporate additional functionality (such as an azide) via nucleophilic displacement of a mesyl group. Ultimately, this led to the incorporation of a triazole of which two analogues were made.

Two additional functionalities were also planned to be introduced, also via the nucleophilic displacement of a mesylate. This included an aniline and a tyrosine derivative. Unfortunately, the former could not be completed due to time constraints and the latter was compounded with synthetic problems.

Unfortunately, none of the compounds synthesised obtained notable activity against Spermidine Synthase.

7. Protein Farnesyltransferase inhibition

For our PFT inhibitor development efforts, we took inspiration from the work of van Voorhis *et al.*¹⁴⁵ This group successfully synthesised and assayed a wide series of inhibitors with low nanomolar potencies *in vivo*. Because of the superior nature of these inhibitors, we based our own designs on this series and expanded the subset.

Derivatisation of indole *N* position – Proof of concept

The inhibitors developed by Voorhis *et al.* were based on a tetrahydroquinoline (THQ) scaffold.¹⁴⁵ The most potent inhibitor in the series, **PB-93** (see Figure 7-1), was co-crystallised with PFT to better understand its mode of action (PDB ID: 2R2L). After analysing the crystal structure, we found several key interactions:

- 1) The THQ scaffold has a pi-pi interaction with a tyrosine residue which is located directly underneath it.
- 2) An imidazole, which extends from the THQ scaffold, forms a chelating interaction with a zinc ion located within the active site.
- 3) A nitrile, which also extends directly from the THQ scaffold, occupies a hydrophobic cavity within the active site.

To develop our own variation of this inhibitor, we substituted the THQ scaffold with an indole. An indole, similar to quinoline, is also a biologically privileged structure meaning that the structure exhibits good pharmacodynamic and pharmacokinetic properties. This novel scaffold allowed us to retain interactions (Figure 7-1). By superimposing this new molecule, our proof of concept, onto **PB-93** within the crystal structure, we could see how this new molecule would be well accommodated within the active site (Figure 7-2). Once we have successfully synthesised this proof of concept compound **58** and demonstrated activity, we would expand the indole at positions 2' and 3' to increase the site occupancy to refine the selectivity and potency.

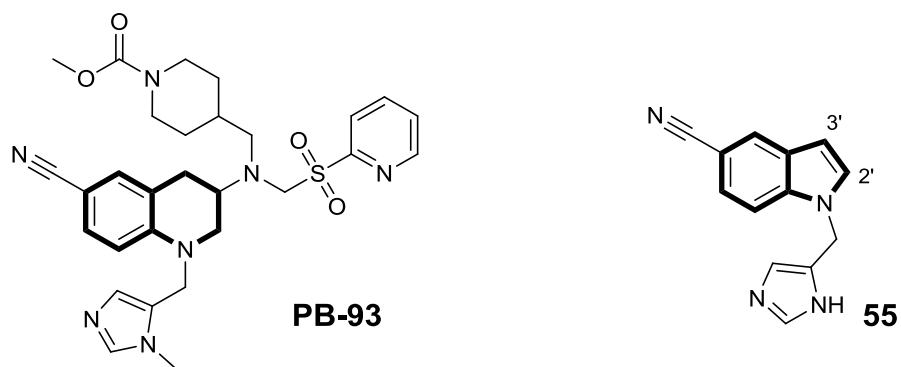


Figure 7-1 Tetrahydroquinoline based inhibitor developed by van Voorhis *et al.* and our indole derivative.¹⁴⁵

The scaffold of the molecules have been bolded to highlight the similarity between them. The indole 3' position would be expanded further to develop the compound series.

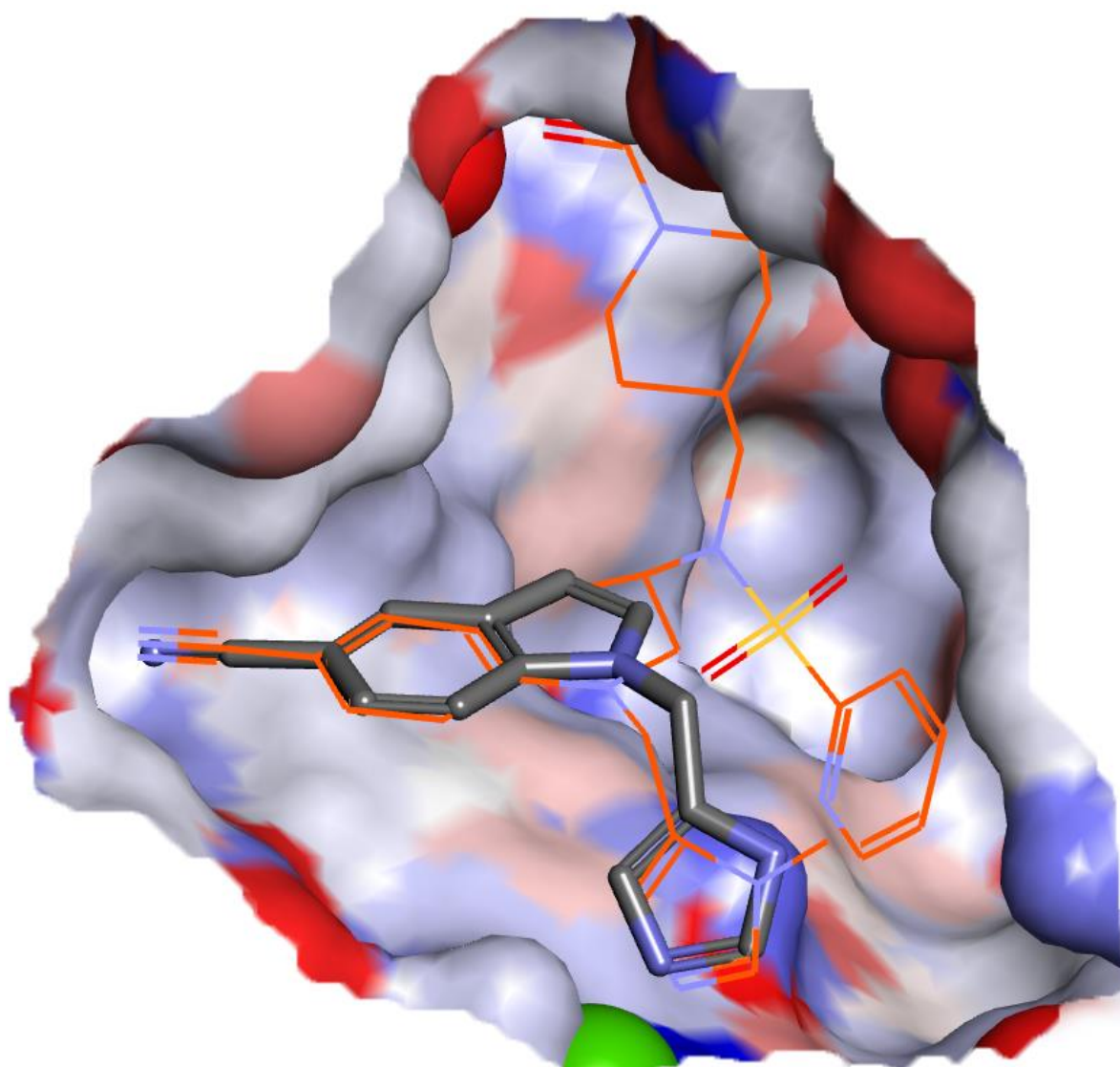
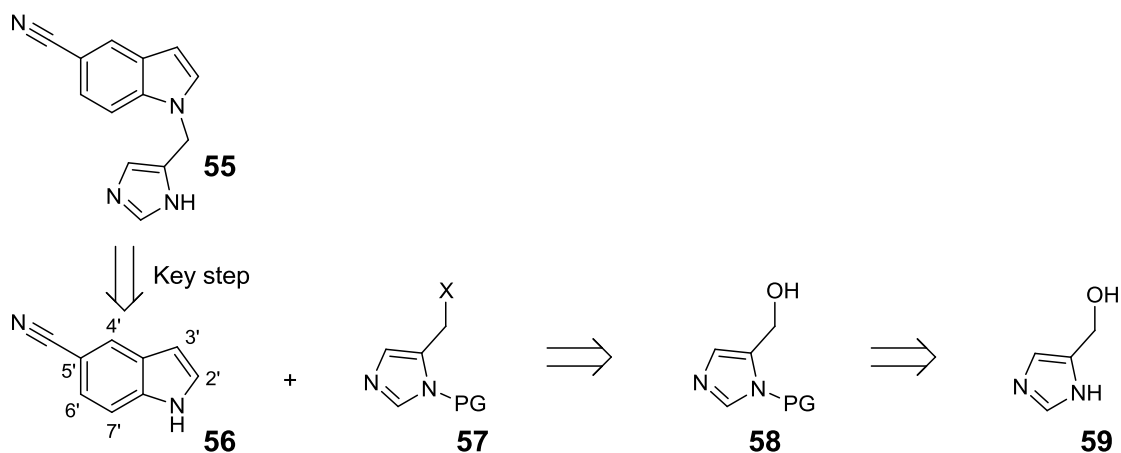


Figure 7-2 Wide entrance to the active site of protein farnesyltransferase (PDB code: 2R2L). Our proof concept indole derivative 55 with a CDOCKER score of 25 is superimposed on the crystallised tetrahydroquinoline scaffold (carbons in orange, CDOCKER score 55). Note the imidazole chelating to the zinc (green) and the nitrile filling the hydrophobic cavity to the left. The groove along the right is the binding site for the farnesyl moiety.

Synthetic planning

To make this proposed proof of concept compound (Figure 7-1), the main step would involve attaching the indole to the imidazole derivative. To achieve this key step (see Scheme 7-1), we would exploit the indole nitrogen's nucleophilicity to attach the imidazole functionality. The imidazole residue, would then be derivatised with a suitable leaving group at the appropriate position to make it electrophilic. To this aim, we sourced an imidazole **59** which was substituted with a hydroxyl group. The hydroxyl would then be converted to a chloride to impart electrophilicity at the appropriate position. Moreover, this imidazole was not methylated like **PB-93**. In order to avoid the imidazole acting as a nucleophile, it would be first need to be protected with a base stable protecting group.

Scheme 7-1 Proposed disconnection of the proof of concept **58**

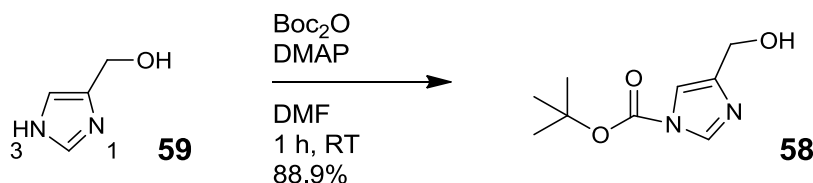


Synthesis

Imidazole derivative

The synthesis of the imidazole zinc chelating moiety half of the proof of concept compound was commenced.

Scheme 7-2 Synthesis of 58



We envisaged that from (1H-imidazol-4-yl)methanol **59**, which is commercially available, the primary alcohol could be converted into a convenient leaving group, facilitating the substitution of this moiety onto the indole. To negate the possibility that the imidazole nitrogen could act as possible competing nucleophile, we used a base stable protecting group on it. Boc anhydride (Boc₂O) fulfilled this criterion. The Boc group was introduced in good yield using a catalytic amount of dimethoxy-aminopyridine (DMAP) as acyl carrier. After following the reaction with TLC, the product **58** was purified using column chromatography. This removed most impurities, including the characteristic Boc protection by-product - *tert*-butanol. The reaction was deemed to be successful with the appearance of the characteristic NMR signal which integrates for 9 protons in the upfield region. It was interesting to note that this signal was divided into two separate peaks: imidazole has two tautomeric states, both being susceptible to Boc protection. Therefore, we concluded that two regioisomers had formed. Judging by the relative ratio of these two peaks, roughly 83% of the Boc attached to the nitrogen at the 3' position, while the 1' nitrogen accepted the remainder (see Figure 7-3). Taking the sterics of the primary alcohol into account, it can be presumed that the more significant amount of the Boc group went on to the 3' amine, whilst the remainder went on the more sterically hindered 1' amine. This presumption could be confirmed by an NOE experiment, but was deemed unnecessary as the Boc would soon be removed without impacting the end result.

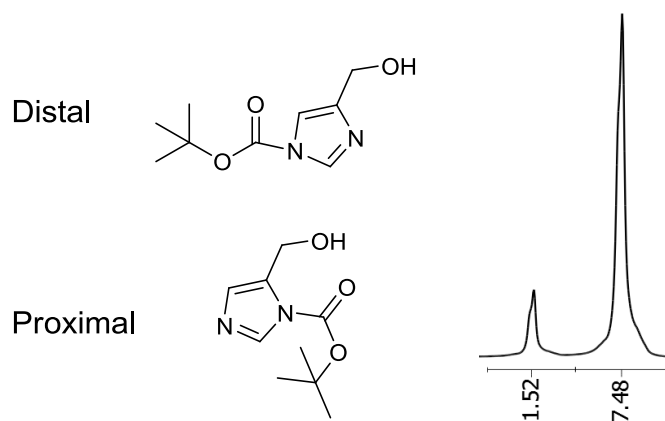
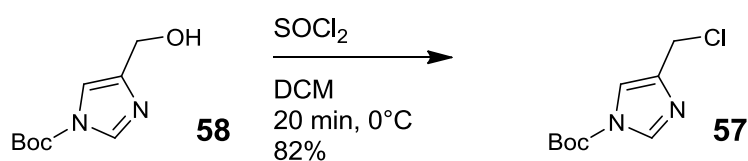


Figure 7-3 NMR segment of the *tert*-butyl group
The signal split is proportional to the locale of the Boc protection group (Integration indicated below).

Scheme 7-3 Synthesis of 57

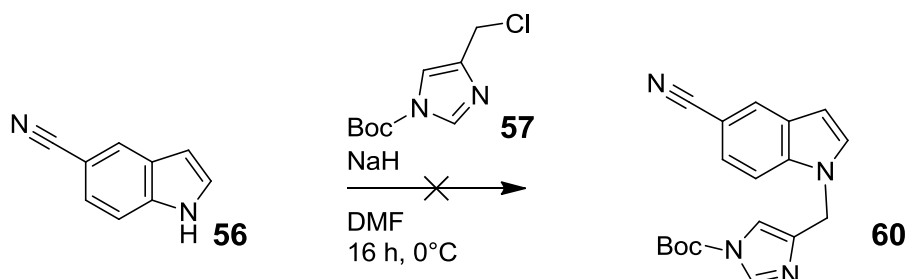


After the successful protection of the amine, the hydroxyl needed to be converted into an appropriate leaving group. For our purposes, we used thionyl chloride for the conversion to the respective chloride **57**.²¹⁷

In this reaction, HCl would be released as a by-product. This posed a potential problem: If the medium turned acidic, the Boc group could inadvertently be lost. To minimise this, the reaction was performed at 0°C to limit the possibility of this unwanted deprotection reaction and was closely followed by TLC. As soon as the starting material had been consumed, the reaction was quenched and the product was isolated. Finally, the product was purified by column chromatography. ¹H NMR spectroscopy confirmed that the chloride was formed owing to the change in chemical shifts and the disappearance of the broad hydroxyl signal.

Key step

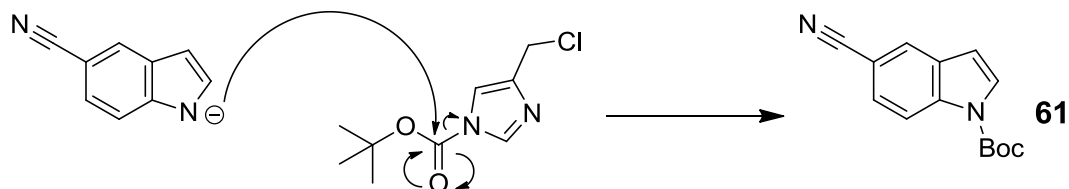
Scheme 7-4 Attempted synthesis of 60



With the Boc-protected electrophile in hand, we were ready to attach it to the indole. The indole nitrogen was deprotonated with sodium hydride to generate the desired nucleophile. After 10 minutes, the evolution of hydrogen ceased and the electrophilic imidazole was

added. The reaction was then left to proceed overnight. After purifying with column chromatography, we discovered an unexpected result: Instead of the nucleophile displacing the chloride, it turned out that the carbamate was the more electronegative functional group and instead resulted in the *N*-Boc-indole **61** (see Scheme 7-5). With this unexpected result, we were back to the drawing board to consider an alternative strategy to this problem.

Scheme 7-5 Mechanism of how the unexpected product, 61, had formed.



An alternative strategy

Close examination of the imidazole functionality in **PB-93** observed that the one nitrogen, opposite to the zinc chelating nitrogen, was methylated.^{145, 218} If we were to revise the proof of concept compound to incorporate this methylated functionality **62**, the previously experienced issues would be resolved. Furthermore, we did not expect that the binding efficacy would be adversely affected as this methyl group faces directly away from the zinc atom (refer to Figure 7-4).

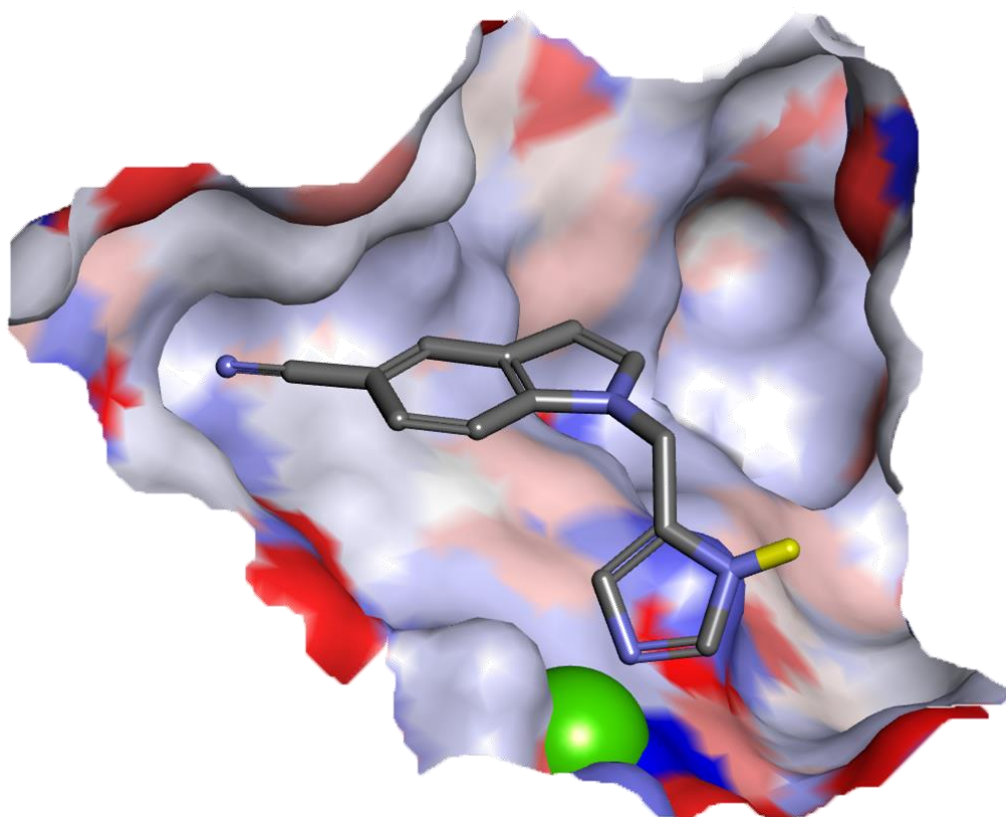
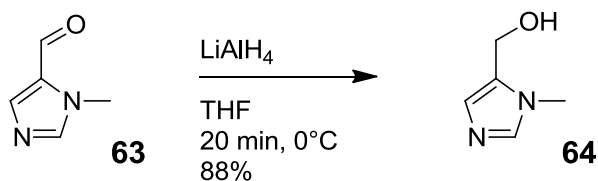


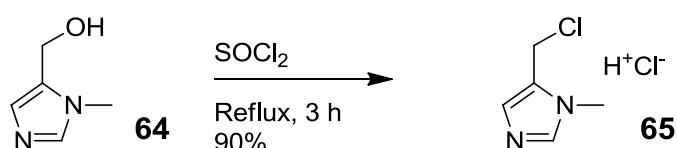
Figure 7-4 Alternative proof of concept compound
The atom highlighted in yellow was to be substituted from a proton to a methyl give **62**, similar to **PB-93**.

Scheme 7-6 Synthesis of 64



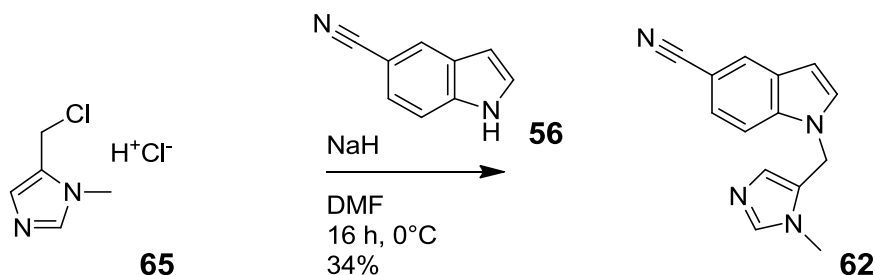
We were able to purchase a suitably methylated imidazole derivative as the aldehyde **69**, which we promptly reduced using LiAlH_4 .²¹⁸ Due to the amphoteric nature of this molecule, excess LiAlH_4 was quenched using $\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$. ^1H NMR spectrum confirmed the aldehyde was successfully reduced owing to the disappearance of the upfield aldehyde singlet and the appearance of a broad hydroxyl signal.

Scheme 7-7 Synthesis of 65



To make the hydroxy-substituted carbon electrophilic once again, the alcohol was converted to the chloride.²¹⁸ In contrast to the previous chlorination where we added a stoichiometric amount of thionyl chloride, a vast excess of SOCl_2 was now added. Under these conditions, the chloride was now isolated as the hydrochloric salt. Trituration with diethyl ether afforded the purified chloride salt. The ^1H NMR spectrum obtained agreed with what we expected, with no evident broad hydroxyl signal. Because the solvent for the NMR experiment was different, we could not make a direct comparison to the chemical shifts of the signals. In retrospect, mass-spectrometry would have given more confidence that that reaction was chloride was obtained, but this was not performed at the time.

Scheme 7-8 Synthesis of 62



With the required precursor now in hand, we were ready for a second attempt to link the major two components together. Once again, sodium hydride was used to deprotonate the indole **59** to produce the nucleophile, using an additional equivalent of sodium hydride to neutralise the hydrochloric acid salt of the imidazolic chloride. Once bubbling ceased, the chloride **71** was added. Because the generated indole anion would act as a base to the acidic crystal salts, we used an extra equivalent sodium hydride to neutralise the salt. The

reaction was left to stir overnight. After workup, the product was purified by chromatography. ^1H NMR spectroscopy indicated that all expected signals were present, all of which integrated appropriately. Unfortunately, not all peaks could be unambiguously assigned due to overlapping signals. It was however enough to confirm that we successfully synthesised the desired proof of concept compound. As the crystal structure was solved, ^{13}C NMR characterisation was not performed. Although only a modest yield was attained, it was sufficient for testing purposes.

The IC_{50} in whole cell assay was measured to be $139\ \mu\text{M}$ according to the procedures by Makler *et al.*²¹⁹ Although this is only modest activity we could now incorporate functional groups to optimise hydrophobicity to pass the cell wall and provide additional reversible binding interaction within the large active site of PFT. To characterise this important intermediate further, we determined and published the crystal structure using X-ray crystallography (refer to Figure 7-5).ⁱ

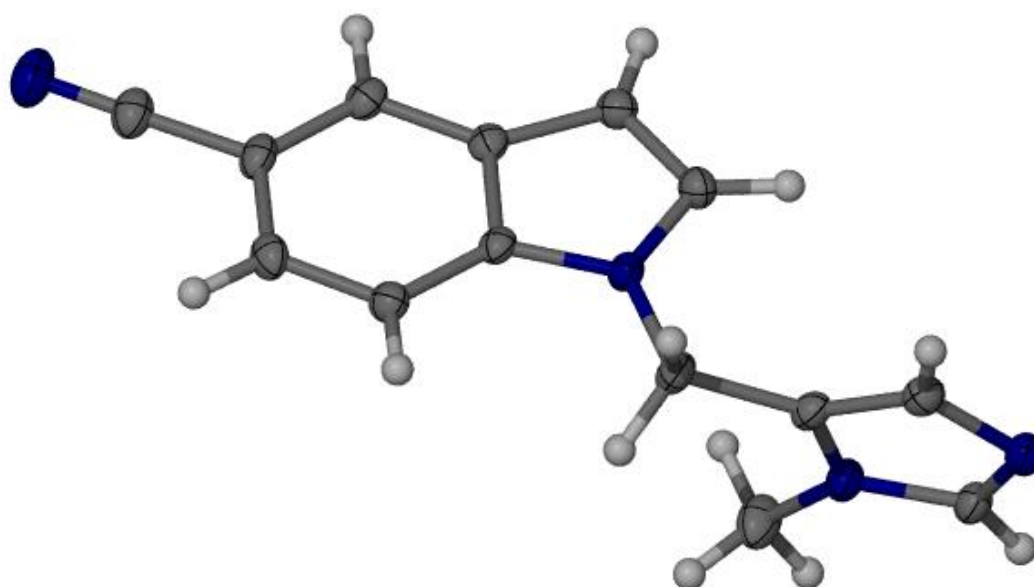


Figure 7-5 Crystal structure of the proof of concept compound, 62.

It is interesting to note that this crystallised form of the ligand is not in the minimum energy conformation. Using (simulated heating) annealing and a subsequent minimisation, the conformation with the lowest potential energy was searched for and compared to the obtained crystal structure. (This method is not suited for entire protein/ligand complexes, but this ligand only has 3 notable rotatable bonds.) When the energies between these two states are compared, a significant difference is observed ($\sim 7\times$).

ⁱ Acta Cryst. (2012). E68, o3486

Table 7-1 Energy differences in crystal structure and the global minimum.

Solvent model	Solid State Energy (kcal/mol)	Global minimum (kcal/mol)
None	147.064	22.5184
DDDE	144.838	20.1710
GBMV	145.643	19.5651
PBSA	145.917	20.7626

This difference is exemplary of how different crystal packing may affect the conformation of even a small molecule. It also highlights the requirement of a minimisation step when preparing protein crystal structures for modelling application, especially if the structure is of low resolution ($> 2 \text{ \AA}$). However, it should be noted that these energy values are dependent on the equilibrium values given in the CHARMM forcefield.

Derivatisation indole 3' position – (Xylene)

Familiarisation of the biochemical environment

As a significant amount of work has already been done in this area, we needed to harness this knowledge to better understand the biochemical environment of the target site.^{143, 145, 220-223} Previous work included the publication of numerous PFT crystal structures of various classes of ligands, including ethylenediamine, tetrahydroquinoline and piperidine based structures. We would attempt to combine some of the best features of these compounds.

To do this, we downloaded all *Rattus norvegicus* PFT crystal structures from the PDB with a resolution of $\leq 2.3 \text{ \AA}$ and superimposed them to the 2R2L crystal structure. In DS, this is done by examining the primary sequence of the proteins and aligning the backbone of the proteins into the same co-ordinate frame. The residues in or near the active site were then examined one by one. The result is shown in Figure 7-6. From this, we expected to see three things: A) the degree of conformational change that could be as a result of the induced fit from the various ligands;²²⁴ B) the key interaction sites of the known inhibitors; and C) the conserved moieties of known inhibitors and their respective locations.

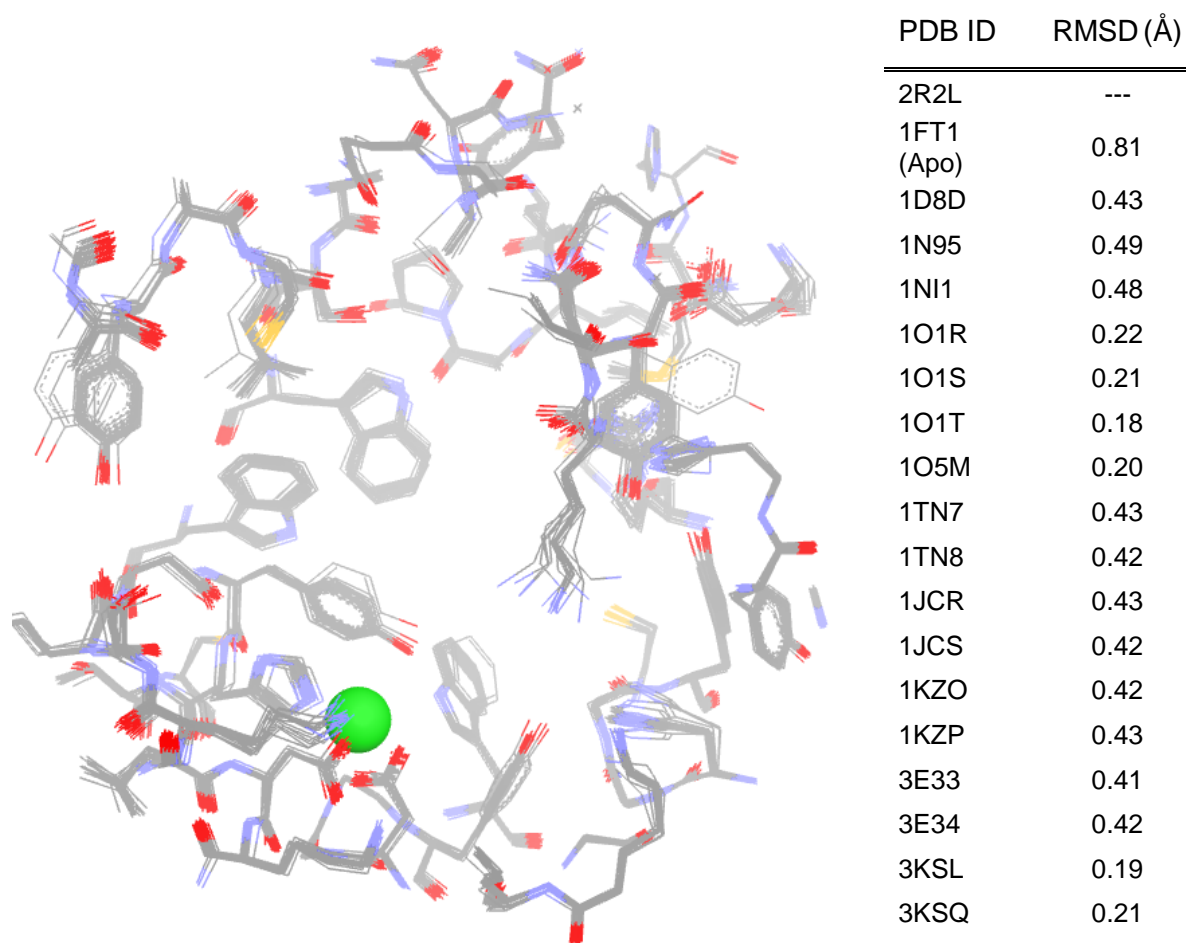


Figure 7-6 Superimposition of the active site of 18 PFT structures, co-crystallised with various ligands.

By referring to the RMSD's of the protein structures relative to 2R2L, we can see that there is minimal deviation in the various crystal structures. Indeed, it is the apo-enzyme crystal structure with the largest deviation. For reference, the active site Zn is shown in green.

It was quickly evident that there was minimal alteration as result of induced fit apart from the phenol rings of two tyrosine residues. The first tyrosine veers slightly left, away from the active site (seen in the upper left in Figure 7-6). The other tyrosine veers strongly away from the active site (seen to the right). None of the three crystal structures with these three induced fit alterations contained ligands which resembled inhibitors which we would design. Therefore, we could safely assume that the structure with PDB ID: 2R2L would be a good choice to use for inhibitor design.

Following this, we examined the ligands to see which interaction sites were significant and whether there were any apparent trends. We quickly noted that several inhibitors contained aromatic rings in similar regions. The enzyme's affinity for aromatic rings is reinforced with the knowledge that the enzyme has strong affinity for certain aromatic amino acid containing protein sequences (the natural substrates). Moreover, the locale of these aromatic amino acids was also easily accessible from the indole 3' position. To exploit this, we would to

attach a benzene ring to this location. Further, in order to reduce the conformational entropy penalty, we would include methyl groups to help keep the molecule in the bioactive shape. Although a single methyl would suffice, the resulting molecule, **63**, would then have the added complexity of being in either of two rotamers (see Figure 7-7). **63** was docked into the prepared 2R2L structure (see Figure 7-8). Although a CDOCKER score of 25 was obtained (which is equal to the indole proof of concept compound), we believed that the xylene would increase the affinity for the active site.

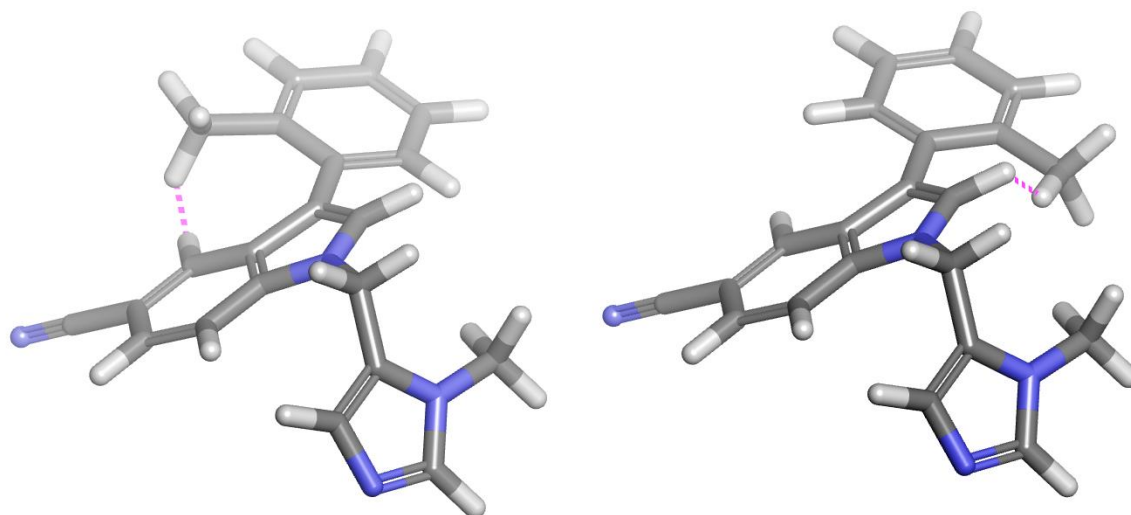


Figure 7-7 Demonstration of possible rotamer creation

Possible steric clash (indicated in purple) of a mono-methylated benzyl group which would result in two rotamers, depending whether the benzylic group is attached with the methyl coordinated to the top or bottom of the indole plane.

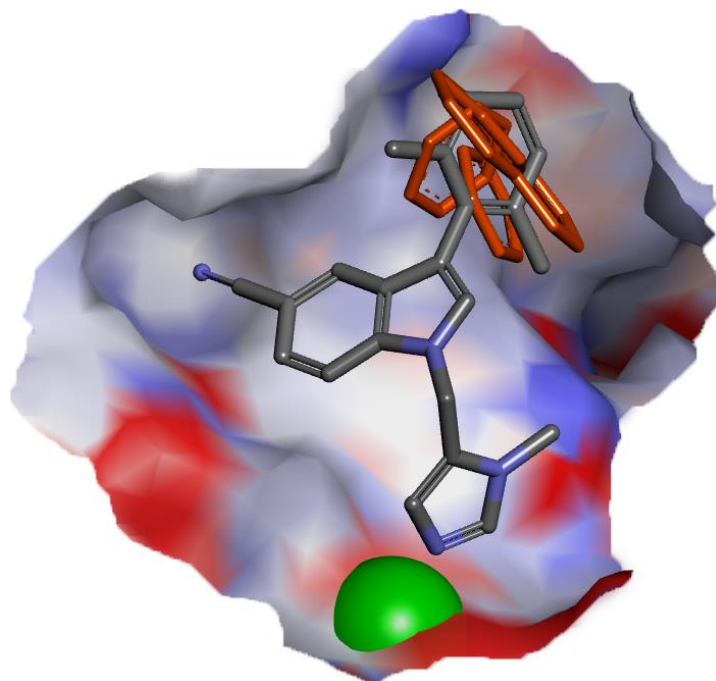


Figure 7-8 Active site surface of 2R2L docked with the planned indole-xylene derivative **63**. The position of the aromatic groups of other inhibitors are also shown in orange while the active site Zn ion is shown in green.

Planning

Albeit with a CDOCKER score of 25, similar to the proof of concept compound **62**, the xylene-coupled indole **63** looked to be a promising compound. A strategy was soon developed to access this molecule efficiently.

Functionalisation at the 3' position in the indole is relatively easily accomplished as it is the most nucleophilic position on the molecule: the indole nitrogen's lone pair can delocalise to conjugate through to the 3' position (Figure 7-9).

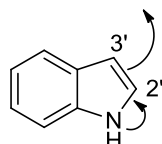


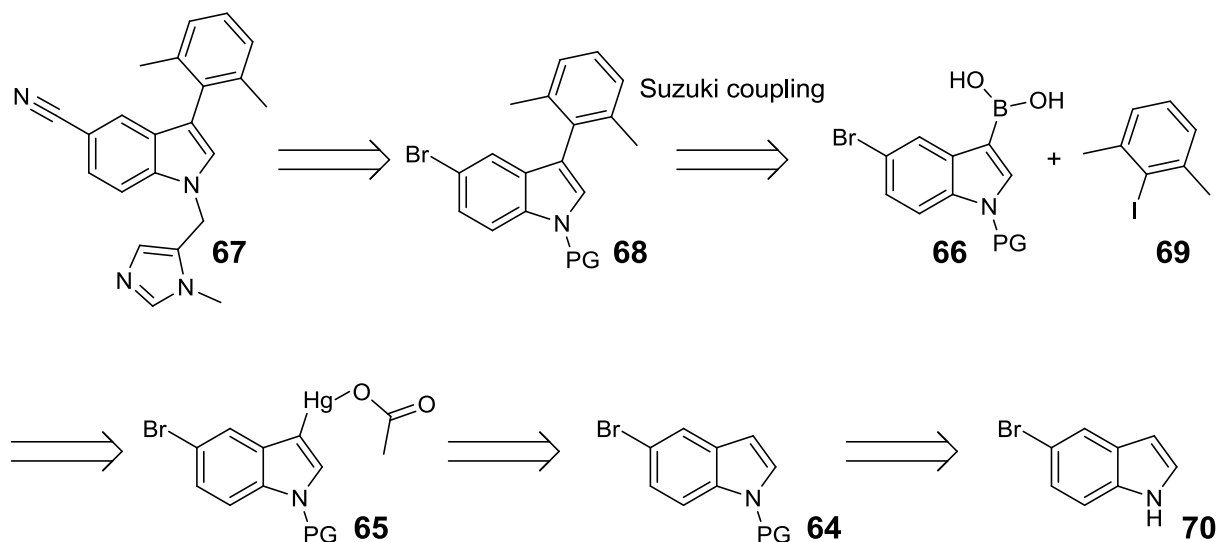
Figure 7-9 Mechanism of indole 3' position's nucleophilicity

To attach the indole directly on to the xylene, we would have to use palladium chemistry, as a conventional substitution reaction would break the xylene ring's aromaticity. To do this, we would employ a Suzuki reaction, which involves oxidative addition and reductive elimination instead.²²⁵ Typically, the conventional method to functionalise the indole 3' position would be to install a halogen (for example, using *N*-bromosuccinimide) and then perform lithium halogen exchange to be able to access a borate ester. From that point, a base work-up would yield the boronic acid from which a Suzuki coupling would be performed to link the xylene to the indole. Unfortunately, we could not use this route because our indole was already substituted with bromide at the 5' position, which would give chemoselectivity issues at the lithium halogen exchange step. We therefore needed an alternative method to access the boronic acid for the palladium coupling reaction.

Xiong *et al* were able to do just this, based on a mercuration reaction first published by Dimrith in 1898 (refer to Scheme 7-9).^{226, 227} Starting from an *N*-protected-5-bromo-indole **64**, the group derivatised the 3' position with mercuric acetate in glacial acetic acid to yield a mercurial compound **65**. This was then converted to the boronic acid **66** using borane. The Suzuki coupling's reagents would then be ready and installation of the nitrile and imidazole moiety would then allow access to target compound **67**.

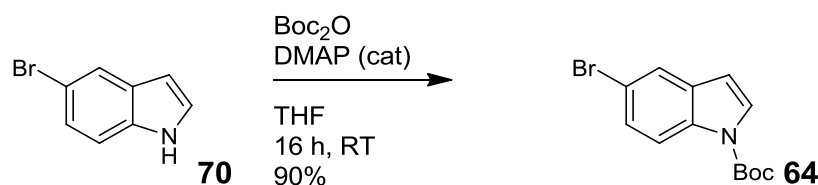
For this synthetic sequence, the group protected the indole nitrogen with a tosylate.²²⁶ Because we had easy access to Boc_2O we used it instead of the tosyl-protection group.

Scheme 7-9 Retrosynthesis of xylene-derivative indole, 63



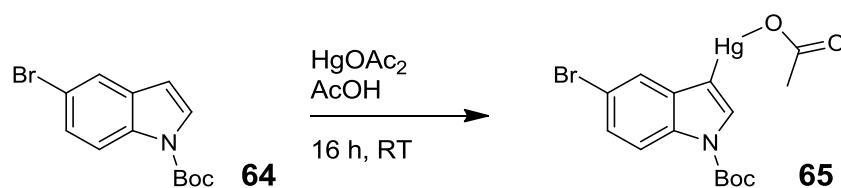
Synthesis

Scheme 7-10 Synthesis of 64



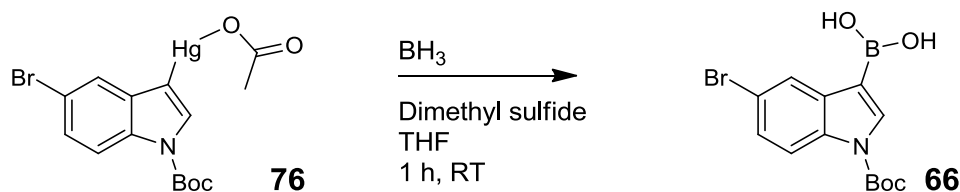
The Boc protected indole was obtained by using Boc anhydride and a catalytic amount of DMAP on 5-bromo-indole **70**.²²⁸ After the reaction was purified by column chromatography in excellent yield, ^1H NMR spectroscopy confirmed the characteristic *tert*-butyl upfield signal which integrates for 9 protons.

Scheme 7-11 Synthesis of 65



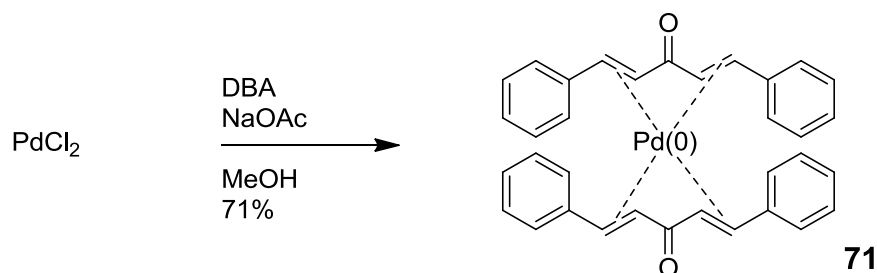
The mercuration was started by dissolving mercuric acetate in glacial acetic acid.²²⁶ To this was added the Boc-protected indole **64**. After the mixture was stirred overnight, a white precipitate was collected by filtration and used without further purification. ^1H NMR spectroscopy then confirmed that the desired mercurial **65** formed due to the disappearance of an aromatic signal and the appearance of an upfield singlet which integrates for 3 protons.

Scheme 7-12 Synthesis of 66



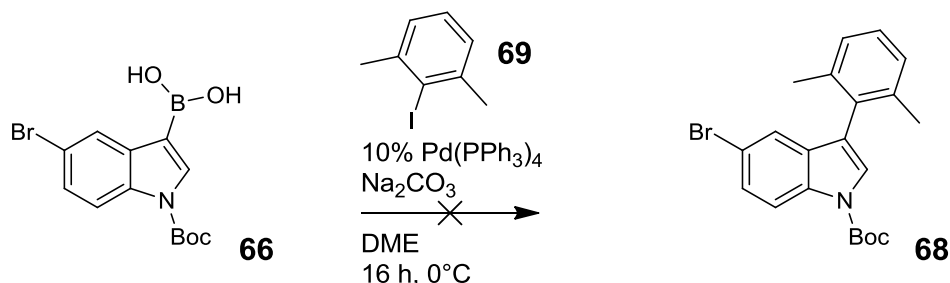
After the isolation of the mercurial **65**, borane was used in dimethyl sulfide to make the boronic acid **66**.²²⁶ Insoluble impurities were removed via filtration and ^1H NMR spectroscopy confirmed the product to be free of significant impurities was then used without further purification. Further, ^1H NMR showed the disappearance of the upfield methyl singlet and the appearance of a broad singlet which integrated for 2 protons. Together with the change in retention factor, this led us to believe that the boronic acid was successfully isolated and our reagent was ready for Suzuki cross-coupling.

Scheme 7-13 Synthesis of 71



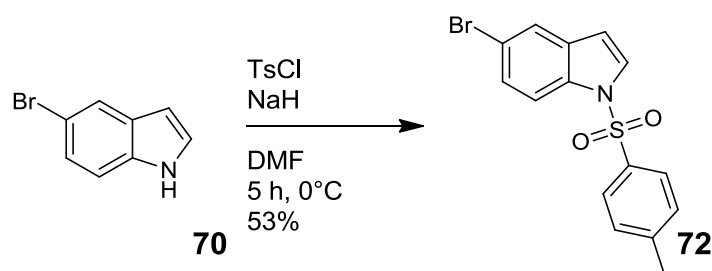
To synthesise the palladium catalyst, we started with a solution of dibenzylideneacetone and sodium acetate in hot methanol.²²⁹ We then added the PdCl_2 and allowed the reaction to stir. After a while, the product started to precipitate, which was furthered by cooling the solution. After filtration with water and acetone, we had pure $\text{Pd}(\text{DBA})_2$ **71**. Although air stable, the reagent was stored shielded away from light in a desiccator. From this reagent, we would prepare tetrakis(triphenylphosphine)palladium(0) (tetrakis) ($\text{Pd}(\text{PPh}_3)_4$) *in situ*.

Scheme 7-14 Attempted synthesis of 68



Finally, with all the reagents in hand, we attempted our Suzuki reaction. Although we were aware that the bromine present on the indole could potentially also be susceptible to derivatisation, we were hopeful that the iodo would be more labile and afford the required chemoselectivity.

As this reaction is very oxygen sensitive, all solutions used were degassed with nitrogen. In a three neck flask, we added Pd(DBA)₂ **71** to dimethoxyethane (DME). This was followed by the addition of PPh₃. The solution soon turned bright yellow which confirmed the successful formation of tetrakis (Pd(PPh₃)₄) *in situ*. This was followed by the addition of the iodo-xylene **69**, boronic acid **66** and Na₂CO₃. The reaction was left to stir overnight, after which it turned green. Upon examination with TLC, numerous by-products formed which we deemed fruitless to purify. Alterations which could be made to make the reaction work would include: the use of a different base (example Cs₂CO₃); preparation of the tetrakis beforehand; employment of freeze thaw techniques to remove all traces of oxygen; variation of the Pd catalyst and the concentration thereof. However, our first action was to eliminate the possibility that it could be the protecting group used. So we changed the indole nitrogen protection from the Boc group to a tosyl, which was used by Xiong *et al.*

Scheme 7-15 Synthesis of 72

Sodium hydride was used to deprotonate the indole nitrogen and afford the nucleophile. To this we added the tosyl chloride.²³⁰ After purification, ¹H NMR spectroscopy confirmed that the tosylation was successful by the appearance of two aromatic doublet signals (both of which integrated for 2 protons) and the appearance of an upfield singlet which integrated for 3 protons.

Unfortunately, this project's allocated time expired at this point and was halted.

Derivatisation indole 3' position – (Maleimide)

In another effort to expand deeper into the binding cavity, we would synthesise an indole **73** that is derivatised with a maleimide at the 3' position (refer to Figure 7-10).

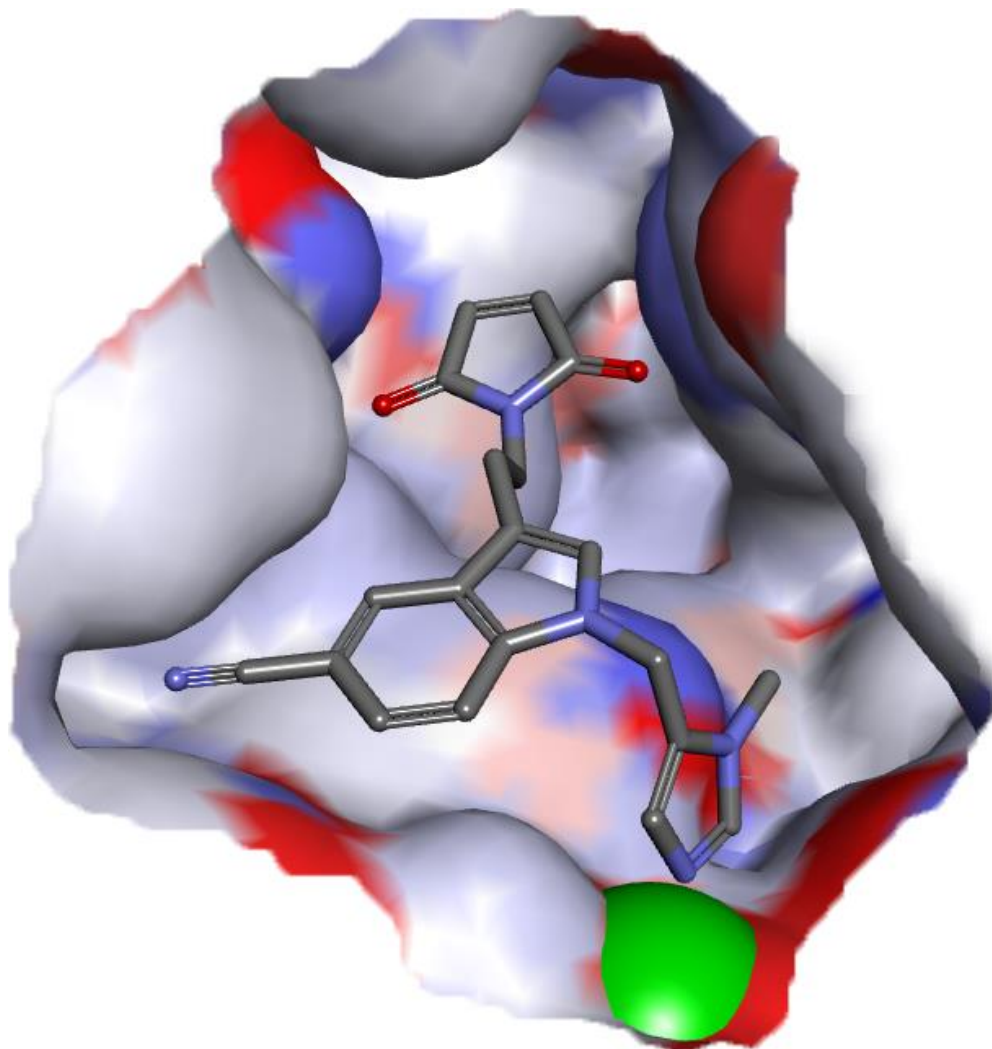
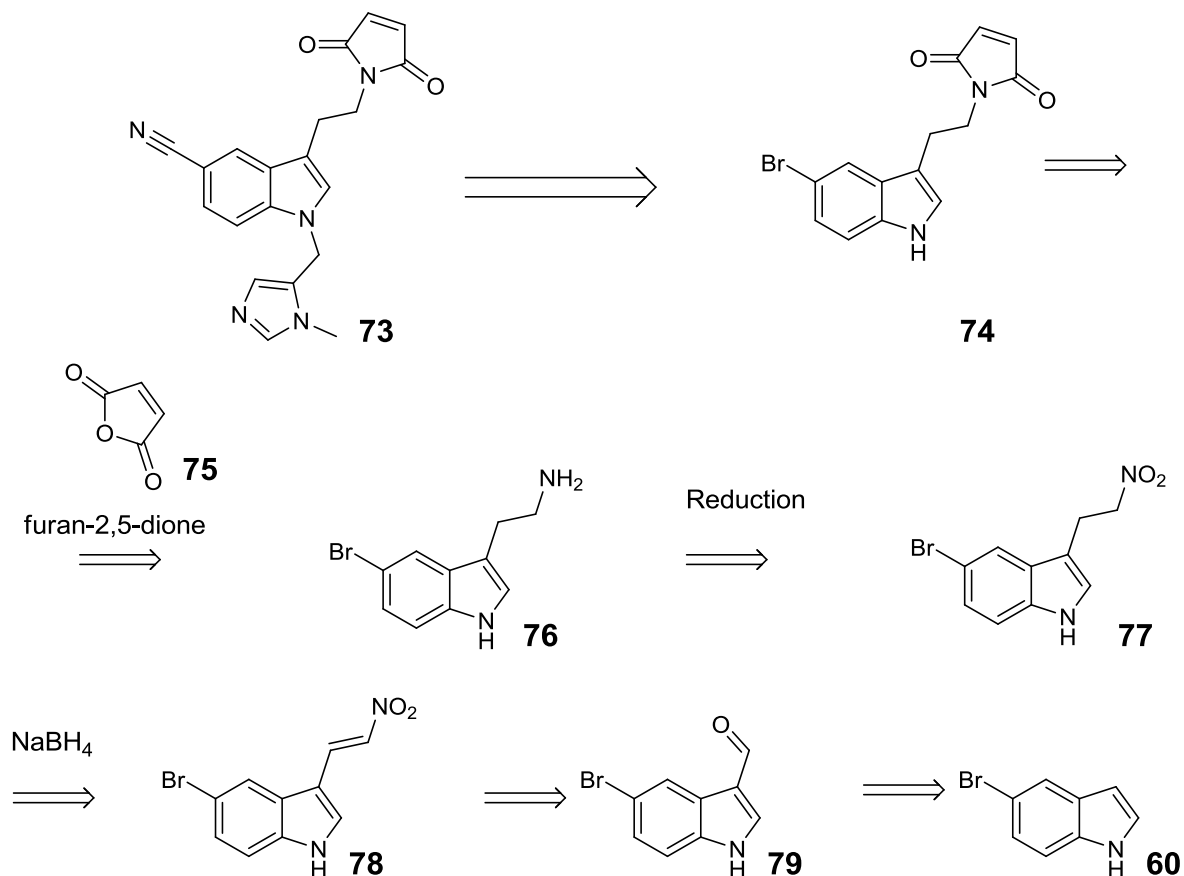


Figure 7-10 Maleimide-indole **73** docked in the active site of PFT. The moiety extends from the indole 3' position. Again, the green sphere indicates the catalytic zinc ion.

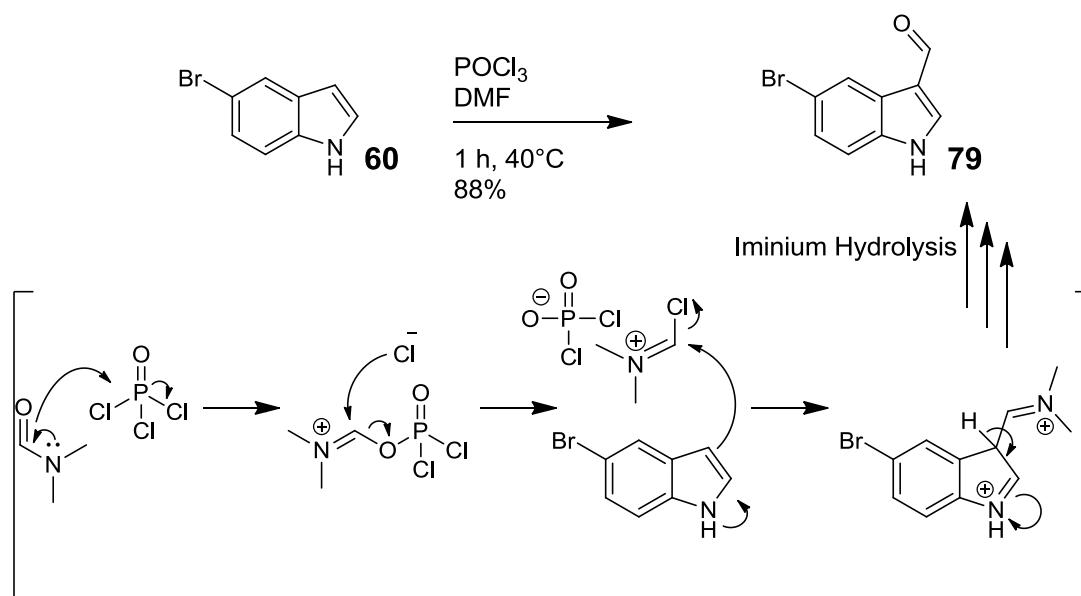
Planning

Once again, the first disconnection points would be between the imidazole and nitrile functionalities to the indole. From here, the maleimide disconnection begins. The maleimide **74** can be accessed from the addition of furan-2,5-dione **75** to the amine **76**. This amine can be obtained from the reduction of the nitro **77**, using a suitable reducing agent like Raney Nickel which is preceded by the reduction of the alkene **78**. The nitro, in turn, could be prepared by the use of a Henry reaction from the aldehyde **79**. Finally, the aldehyde can be constructed by exploiting the nucleophilicity of the 3' position in the indole **70** using a Vilsmeier-Haack reaction.²³¹

Chapter 7 – Protein Farnesyltransferase inhibition

Scheme 7-16 Disconnection of indole-maleimide **73**

Synthesis

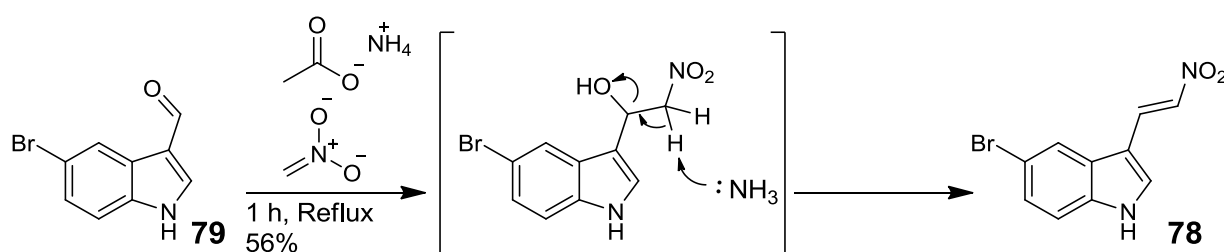
Scheme 7-17 Synthesis of **80**²³²

We started our synthesis of the aldehyde **79** using a Vilsmeier-Haack formylation based on precedence set by Schumacher *et al.*²³² This started with the dropwise addition of

Chapter 7 – Protein Farnesyltransferase inhibition

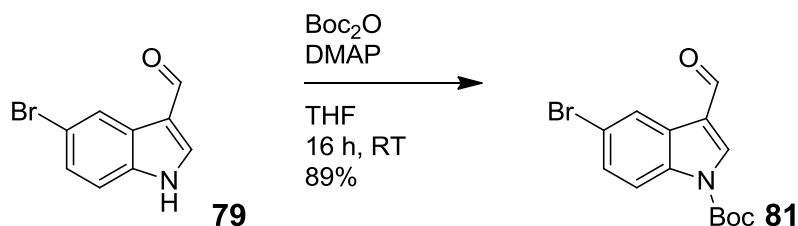
phosphoryl chloride to DMF. Once the electrophilic Vilsmeier salt was formed, the nucleophilic indole was added to form the iminium ion. Finally, aqueous base was added to hydrolyse the iminium ion to the aldehyde. Filtration was sufficient to yield pure product in very good yield.

Together with an increased polarity (determined by TLC), ^1H NMR spectroscopy confirmed that the correct product had formed: The spectrum indicated the presence of a far downfield singlet which integrated for 1, consistent with the presence of an aldehyde proton. Also, the remainder of the signals were unambiguously assigned.

Scheme 7-18 Synthesis of **78**²³³

The installation of the required nitro group by a Henry reaction was performed according to the procedure of Muratore *et al.*²³³ Under reflux conditions, nucleophilic nitromethane attacked the aldehyde to form a hydroxyl intermediate. The alcohol is then eliminated in an E1cB mechanism to produce the double bond.²³³ With the aid of TLC, it was evident that product formation occurred, but during the work-up water wash step, the product degraded. Subsequent attempts at this reaction skipped this step to directly purify the product with column chromatography. Unfortunately, poor solubility compounded these efforts and only a modest yield was attained.

The bright orange product was analysed by ^1H NMR spectroscopy which led us to believe that we had obtained the expected nitro **78** owing to the disappearance of the aldehyde singlet and the appearance of two signals which integrate for 1 proton each.

Scheme 7-19 Synthesis of **80**

The next step in the sequence was to reduce the double bond of the nitro-indole derivative **78**, but this was compounded by solubility issues. In an attempt to alleviate the solubility

problem, we would increase the molecule's hydrophobic character by introducing Boc protection on the indole nitrogen.

In a standard Boc protection method, Boc-anhydride was added to the indole in THF and a catalytic amount of the acyl transfer agent, DMAP, was added. After overnight stirring, the reaction was worked up and purified. ¹H NMR spectroscopy indicated the tell-tale upfield singlet which integrated for 9 protons. All other signals present were accounted for, which convinced us that the carbamate **80** had indeed formed in excellent yield.

Unfortunately, due to the fact that the project focus required a shift in the direction of the SpdSyn inhibitors (the required assay would only be available for a limited period of time), further work on this synthetic route was halted.

Conclusion

Synthesis of the proof of concept compound **62** was successfully completed and was demonstrated to inhibit parasitic growth in whole cell assay. However, it still remains to be demonstrated, using an enzymatic assay method, that the target is indeed the enzyme Protein Farnesyltransferase.

Future work in this project should include functionalising the indole at the 3' position to allow for increased occupancy of the enzymatic cavity. This should include additional attempts to introduce an aromatic moiety, using modified methods of palladium chemistry.

8. Experimental

Synthetic Procedures - General

Solvent purification

Reaction solvents were purified by distillation under nitrogen atmosphere using the following drying agents:²⁰²

THF	Sodium flakes and benzophenone as indicator
Diethyl ether	Sodium wire
Toluene	Sodium chunks
DCM	CaH ₂
Methanol	CaH ₂ with a catalytic amount of iodine
DME	Used straight from anhydrous container
Pyridine	Potassium hydroxide

Chromatography

Thin layer chromatography used Alugram XTRA Silica G/UV254 plates. Visualisation of the analytes was achieved by: UV fluorescence, iodine, vanillin, KMnO₄ or ninhydrin staining.

Column chromatography was employed for sample purification. Gravity chromatography used 400 mesh silica, whilst flash silica used 60 mesh size whilst under pressure from compressed air using generally accepted procedures.²³⁴

Spectroscopy

NMR: Spectra were performed on Varian INOVA spectrometers (300 and 400 MHz) from the Central Analytical Facility from University Stellenbosch. In all cases, residual solvent signals were used as reference for chemical shifts, reported in ppm (CDCl₃ = 7.260, DMSO-d₆ = 2.500, MeOD-d₄ = 3.310).

X-ray diffraction: Bruker APEXII CCD diffractometer

Spectrometry

MS: Waters API Q-TOF Ultima spectrometer, using electron spray ionisation.

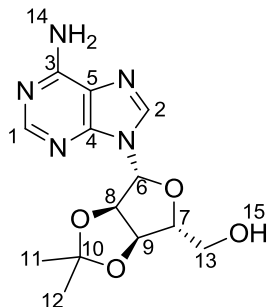
General

Reactions were performed under inert atmospheres (N₂), unless otherwise stated. Glassware was oven or flame dried. Reaction containers were purged with nitrogen which had been passed through a column of 4 Å molecular sieves.

In vacuo solvent removal entails the use of a rotary evaporator to remove solvent. When required, a high vacuum pump was used to remove the remaining traces of solvent under ~ 0.1 mm Hg.

Synthetic Procedures (Chapter 5)

3 ((3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methanol

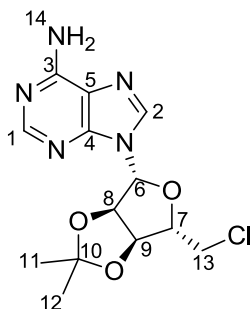


To an anhydrous solution of adenosine **1** (500 mg, 1.87 mmol) in acetone (100 mL) was added PTSA (3.20 g, 18.7 mmol) and molecular sieves. The reaction mixture was allowed to stir for 24 hours under an anhydrous nitrogen atmosphere. The solution was made basic by the addition of ice cold, saturated NaHCO₃ solution. The product was extracted with DCM (3x 150 mL), washed with brine and dried with MgSO₄. After removal of the solvent *in vacuo*, the solid was purified by chromatography (1:9 MeOH:DCM) to afford a white product (63 mg, 11%). *R_f* = 0.47 (1:9 MeOH:DCM). ¹H NMR (400 MHz, DMSO) δ 8.34 (s, 1H, **1 or 2**), 8.15 (s, 1H, **1 or 2**), 7.34 (s, 2H, **14**), 6.12 (d, *J* = 3.1 Hz, 1H, **6**), 5.34 (dd, *J* = 6.2, 3.1 Hz, 1H, **8**), 5.24 (t, *J* = 5.6 Hz, 1H, **15**), 4.96 (dd, *J* = 6.2, 2.5 Hz, 1H, **9**), 4.21 (td, *J* = 4.8, 2.5 Hz, 1H, **7**), 3.63 – 3.45 (m, 2H, **13**), 1.54 (s, 3H, **11 or 12**), 1.32 (s, 3H, **11 or 12**).

3 (Alternative synthesis)

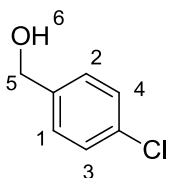
To an anhydrous solution of adenosine **1** (2500 mg, 9.355 mmol) in acetone (60 mL) was added PTSA (19.33 g, 112.3 mmol), dimethoxypropane (3897 mg, 37.42 mmol) and molecular sieves. The reaction mixture was allowed to stir for 48 hours under an anhydrous nitrogen atmosphere. The solution was made basic by the addition of ice cold, saturated NaHCO₃ solution. The product was extracted with DCM (5x 50 mL), washed with brine and dried with MgSO₄. After removal of the solvent *in vacuo*, the solid was purified by chromatography (1:9 MeOH:DCM) to afford a white powder (2.386 g, 83%). *R_f* = 0.47 (1:9 MeOH:DCM). ¹H NMR (300 MHz, DMSO) δ 8.34 (s, 1H, **1 or 2**), 8.15 (s, 1H, **1 or 2**), 7.35 (s, 2H, **14**), 6.12 (d, *J* = 3.1 Hz, 1H, **6**), 5.34 (dd, *J* = 6.2, 3.1 Hz, 1H, **8**), 5.23 (t, *J* = 5.5 Hz, 1H, **15**), 4.96 (dd, *J* = 6.2, 2.5 Hz, 1H, **9**), 4.21 (td, *J* = 4.8, 2.5 Hz, 1H, **7**), 3.63 – 3.44 (m, 2H, **13**), 1.54 (s, 3H, **11 or 12**), 1.32 (s, 3H, **11 or 12**).

4 9-((3a*R*,4*R*,6*S*,6a*S*)-6-(Chloromethyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine

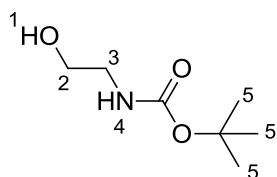


To an ice cold stirring solution of **3** (300 mg, 0.978 mmol) in MeCN (10 mL) was added SOCl₂ (232 mg, 1.95 mmol) and pyridine (255 mg, 3.2 mmol). The reaction mixture was allowed to stir for 16 hours under anhydrous nitrogen atmosphere. The volatiles were then removed *in vacuo* and the crude was dissolved in EtOAc (5 x 50 mL), washed with brine and dried with MgSO₄. Purification with flash chromatography (97.5:2.5 EtOAc:MeOH) yielded a darkish yellow powder (200 mg, 60%). R_f = 0.10 (98:2 EtOAc : MeOH). ¹H NMR (400 MHz, D₂O) δ 8.33 (s, 1H, **1 or 2**), 8.18 (s, 1H, **1 or 2**), 7.37 (s, 2H, **14**), 6.23 (d, *J* = 2.4 Hz, 1H, **6**), 5.48 (dd, *J* = 6.3, 2.4 Hz, 1H, **8**), 5.07 (dd, *J* = 6.3, 2.9 Hz, 1H, **9**), 4.34 (td, *J* = 6.6, 2.9 Hz, 1H, **7**), 3.86 (dd, *J* = 11.2, 7.0 Hz, 1H, **13a**), 3.75 (dd, *J* = 11.2, 6.2 Hz, 1H, **13b**), 1.54 (s, 3H, **11 or 12**), 1.34 (s, 3H, **11 or 12**).

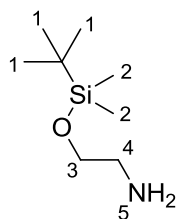
6 (4-Chlorophenyl)methanol



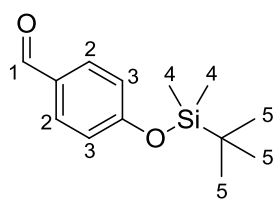
To an ice cold solution of 4-chlorobenzaldehyde **5** (300 mg, 2.13 mmol) in DCM (5 mL) was added NaBH₄ (96 mg, 2.56 mmol) and EtOH (10 mL). After 15 minutes, the reaction was quenched with saturated aqueous NH₄Cl. The DCM and EtOH were removed *in vacuo* whereafter the aqueous phase was extracted with EtOAc (3x 50 mL). The combined organic layers were washed with brine and dried with MgSO₄. After filtration, the solvent was removed *in vacuo* and the crude purified using column chromatography (1:9 EtOAc:Hexanes) to yield an off-white solid (298 mg, 98%). R_f = 0.25 (1:9 EtOAc:Hexanes). ¹H NMR (300 MHz, CDCl₃) δ 7.28 – 7.17 (m, 4H, **1 - 4**), 4.59 (s, 2H, **5**), 1.76 (s, 1H, **6**).

8 Tert-butyl (2-hydroxyethyl)carbamate

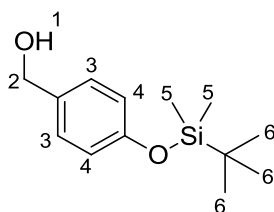
To a stirred solution of ethanolamine **7** (500 mg, 8.19 mmol) in THF (10 ml) was added di-*tert*-butyl dicarbonate (1956 mg, 9.00 mmol). After the addition of a catalytic amount of DMAP, the reaction was left to stir for 16 hours. Solvent was removed *in vacuo* and the remainder was partitioned against EtOAc (3 x 50 ml) and saturated NaHCO₃. The recombined organic layers were then washed with brine and dried with MgSO₄. After filtration, the solvent was removed *in vacuo* and flash chromatography (25:75 EtOAc:Hexanes to 1:1 EtOAc:Hexanes) yielded a colourless oil (553 mg, 84%). $R_f = 0.35$ (1:1 EtOAc:Hexanes). ¹H NMR (300 MHz, CDCl₃) δ 5.02 (s, 1H, **1 or 4**), 3.74 – 3.62 (m, 2H, **2 or 3**), 3.33 – 3.20 (m, 2H, **2 or 3**), 2.49 (s, 1H, **1 or 4**), 1.43 (s, 9H, **5**).

9 2-((Tert-butyldimethylsilyl)oxy)ethanamine

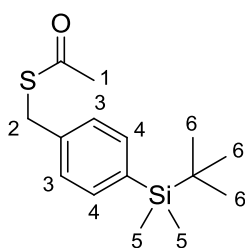
To a stirred solution of ethanolamine **7** (500 mg, 8.19 mmol) and imidazole (669 mg, 9.82 mmol) in DCM (8 ml) was added TBDMSCl (1480 mg, 9.82 mmol). After 1 hour, the reaction was quenched by the addition of water. The solution was extracted with EtOAc (3 x 100 ml) and washed briskly with ice cold 0.1 M HCl. The aqueous phase was extracted with EtOAc (3 x 100 ml). The recombined organic layers were washed with saturated NaHCO₃ and again the aqueous layer was extracted with EtOAc (3 x 100 ml). The recombined organic layers were washed with brine and dried with MgSO₄. After removal of the solvent *in vacuo*, the crude was purified by flash chromatography (2:2:96 MeOH:TEA:EtOAc) to yield a colourless oil (452 mg, 31%). $R_f = 0.32$ (2:2:96 MeOH:TEA:EtOAc) which was visualised with ninhydrin and KMnO₄. ¹H NMR (300 MHz, CDCl₃) δ 3.60 (t, $J = 5.3$ Hz, 2H, **3 or 4**), 2.75 (t, $J = 5.3$ Hz, 2H, **3 or 4**), 1.47 (s, 2H, **5**), 0.88 (s, 9H, **1**), 0.04 (s, 6H, **2**).

11 4-((*Tert*-butyldimethylsilyl)oxy)benzaldehyde

To a stirred solution of 4-hydroxybenzaldehyde **10** (200 mg, 1.64 mmol) in THF (5 mL) was added TBDMSCl (296 mg, 1.97 mmol) and imidazole (167 mg, 2.46 mmol). After 16 hours, the solvent was removed *in vacuo* and partitioned against water and EtOAc (3 x 50 mL). The recombined organic layers were washed with brine, dried with MgSO₄ and filtered. The crude was purified by column chromatography (3:97 EtOAc:Hexanes) to yield a clear oil (304 mg, 80 %). *R*_f = 0.5 (1:9 EtOAc:Hexanes). ¹H NMR (400 MHz, D₂O) δ 9.88 (s, 1H, **1**), 7.86 – 7.81 (m, 2H, **2 or 3**), 7.06 – 7.02 (m, 2H, **3 or 2**), 0.96 (s, 9H, **5**), 0.24 (s, 6H, **4**).

12 (4-((*Tert*-butyldimethylsilyl)oxy)phenyl)methanol

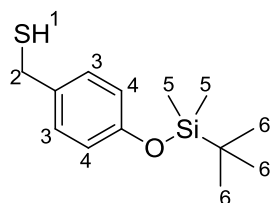
To an ice cold solution of **11** (294 mg, 1.24 mmol) in DCM (5 mL) was added NaBH₄ (57 mg, 1.51 mmol). After 4 hours, the reaction was quenched with a saturated solution of NH₄Cl. The solution was extracted with DCM (3 x 50 mL). The recombined organic layers were washed with brine, dried with MgSO₄, and filtered. Purification with column chromatography (13:87 EtOAc:Hexanes) yielded a clear oil (264 mg, 89%). *R*_f = 0.2 (1:9 EtOAc:Hexanes). ¹H NMR (400 MHz, DMSO) δ 7.21 – 7.16 (m, 2H, **3 or 4**), 6.81 – 6.76 (m, 2H, **3 or 4**), 5.04 (t, *J* = 5.7 Hz, 1H, **1**), 4.41 (d, *J* = 5.7 Hz, 2H, **2**), 0.94 (s, 9H, **6**), 0.17 (s, 6H, **5**).

13 S-4-(*Tert*-butyldimethylsilyl)benzyl ethanethioate

To an ice cold solution of PPh₃ (265 mg, 1.01 mmol) in THF (5 mL) was added D_tBAD (232 mg, 1.01 mmol). After 30 minutes, **12** (150 mg, 0.674 mmol) and thioacetic acid (77.0 mg,

1.011 mmol), dissolved in THF (2 mL), were added dropwise to stir for 2 hours at 0 °C. After the solvent was removed *in vacuo*, the crude was purified by column chromatography (5:95 EtOAc:Hexanes) to yield a colourless oil. $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 7.17 – 7.10 (m, 2H, **3 or 4**), 6.78 – 6.72 (m, 2H, **3 or 4**), 4.07 (s, 2H, **2**), 2.34 (s, 3H, **1**), 0.98 (s, 9H, **6**), 0.19 (s, 6H, **5**).

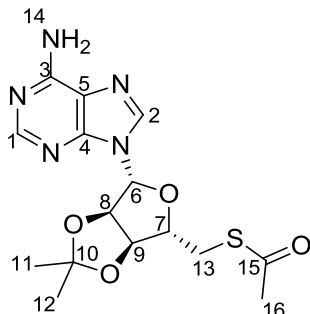
14 (4-((*Tert*-butyldimethylsilyl)oxy)phenyl)methanethiol



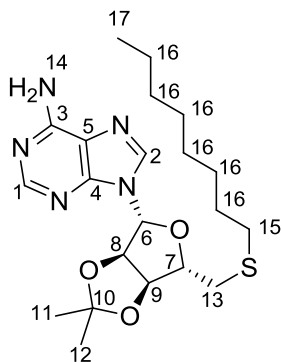
To an ice cold solution of **13** (100 mg, 0.357 mmol) in THF (30 mL) was added LiAlH_4 (17.9 mg, 0.499 mmol). The reaction was closely monitored by TLC and quenched with water. Volatile solvent was removed *in vacuo* and the aqueous layer was extracted with EtOAc (50 mL). The organic layer was washed with brine and dried using MgSO_4 . After the solvent was removed *in vacuo*, the crude was purified by column chromatography (5:95 EtOAc:Hexanes) to yield a colourless oil (57 mg, 20% over 2 steps). $R_f = 0.8$ (1:9 EtOAc:Hexanes). $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 7.17 (d, $J = 8.7$ Hz, 2H, **3 or 4**), 6.77 (d, $J = 8.7$ Hz, 2H, **3 or 4**), 3.70 (d, $J = 7.3$ Hz, 2H, **2**), 1.73 (t, $J = 7.3$ Hz, 1H, **1**), 1.01 – 0.95 (m, $J = 1.6$ Hz, 9H, **6**), 0.22 – 0.15 (s, 6H, **5**).

Synthetic Procedures (Chapter 6)

20 *S*-(((3*aS*,4*S*,6*R*,6*aR*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl) ethanethioate

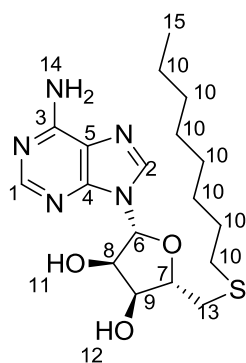


To an ice cold, stirred solution of PPh_3 (950 mg, 5.00 mmol) in THF (5 mL) was added DtBAD (828.9 mg, 5.00 mmol). After the creamy white suspension was stirred for 30 minutes, **3** (500 mg, 1.627 mmol) was added and allowed to stir for an additional 10 minutes at 0°C . Following this, thioacetic acid (266 μL , 3.6 mmol) was added dropwise to turn solution completely miscible and black. The reaction proceeded for 16 hours, after which the solvent was removed *in vacuo*. The precipitate was purified by flash chromatography (8:2 CHCl_3/THF to 9:10 $\text{CHCl}_3:\text{MeOH}$) to yield a yellow foam (590 mg, 98%). $R_f = 0.56$ (1:9 $\text{MeOH}:\text{DCM}$). MS Elemental: $\text{C}_{15}\text{H}_{20}\text{N}_5\text{O}_4\text{S}$. $[\text{H}^+]$: Calc 366.1231; Found 366.1236. ^1H NMR (300 MHz, CDCl_3) δ 8.34 (s, 1H, **1 or 2**), 7.89 (s, 1H, **1 or 2**), 6.12 (s, 2H, **14**), 6.06 (d, $J = 2.1$ Hz, 1H, **6**), 5.50 (dd, $J = 6.4, 2.1$ Hz, 1H, **8**), 4.96 (dd, $J = 6.4, 3.1$ Hz, 1H, **9**), 4.33 (td, $J = 6.9, 3.1$ Hz, 1H, **7**), 3.28 (dd, $J = 13.8, 7.2$ Hz, 1H, **13a**), 3.17 (dd, $J = 13.8, 6.7$ Hz, 1H, **13b**), 2.33 (s, 3H, **16**), 1.58 (s, 3H, **11 or 12**), 1.37 (s, 3H, **11 or 12**). ^{13}C NMR (75 MHz, CDCl_3) δ 194.5 (s, **15**), 155.8 (s, **3, 4 or 5**), 153.2 (s, **1 or 2**), 149.2 (s, **3, 4 or 5**), 139.9 (s, **1 or 2**), 120.3 (s, **3, 4 or 5**), 114.5 (s, **10**), 90.9 (s, **6**), 86.1 (s, **7**), 84.2 (s, **8**), 83.7 (s, **9**), 31.3 (s, **13**), 30.6 (s, **16**), 27.1 (s, **11 or 12**), 25.4 (s, **11 or 12**).

21 9-((3a*R*,4*R*,6*S*,6a*S*)-2,2-Dimethyl-6-((octylthio)methyl)tetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine

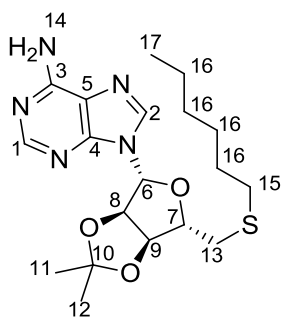
To a solution, cooled to -15 °C, of **20** (50 mg, 0.137 mmol) in freshly distilled MeOH (5 mL) was added NaOMe in MeOH 25% m/v (30 µL, 0.137 mmol). After 45 minutes, 2-bromooctane (132 mg, 0.685 mmol) was added and allowed to stir for 24 hours. The reaction was quenched with water and the solvent removed. After the crude was lyophilised overnight, flash chromatography (1:99 MeOH:DCM to 4:96 MeOH:DCM) afforded an oil (18 mg, 30.2%). $R_f = 0.56$ (1:9 MeOH:DCM). MS Elemental: $C_{21}H_{34}N_5O_3S$. $[H^+]$: Calc 436.2377; Found 436.2382. 1H NMR (300 MHz, $CDCl_3$) δ 8.34 (s, 1H, **1 or 2**), 7.92 (s, 1H, **1 or 2**), 6.08 (d, $J = 2.2$ Hz, 1H, **6**), 5.91 (s, 2H, **14**), 5.51 (dd, $J = 6.4, 2.2$ Hz, 1H, **8**), 5.05 (dd, $J = 6.4, 3.1$ Hz, 1H, **9**), 4.39 (ddd, $J = 7.5, 6.3, 3.1$ Hz, 1H, **7**), 2.83 (dd, $J = 13.6, 7.5$ Hz, 1H, **13a**), 2.75 (dd, $J = 13.6, 6.3$ Hz, 1H, **13b**), 2.52 – 2.45 (m, 2H, **15**), 1.64 – 1.13 (m, 18H, **11, 12 and 16**), 0.89 – 0.81 (m, 3H, **17**). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.6 (s), 153.1 (s, **1 or 2**), 149.3 (s), 140.0 (s), 120.3 (s), 114.4 (s), 90.9 (s, **6**), 86.8 (s, **7**), 84.1 (s, **8**), 83.8 (s, **9**), 34.4 (s, **13**), 32.7 (s, **15**), 31.8 (s, **16**), 29.6 (s, **16**), 29.2 (s, **16**), 29.1 (s, **16**), 28.8 (s, **16**), 27.1 (s, **11 or 12**), 25.4 (s, **11 or 12**), 22.6 (s, **16**), 14.1 (s, **17**).

22 (2R,3R,4S,5S)-2-(6-Amino-9H-purin-9-yl)-5-((octylthio)methyl)tetrahydrofuran-3,4-diol



To a solution of **21** (12.94 mg, 0.0297 mmol) in water (7 mL) was added formic acid (2 mL). The reaction was left to stir for 4 days. Volatiles were removed *in vacuo* to yield pure white powder (11.75 mg 100%). $R_f = 0.63$ (1:9 MeOH:DCM). MS $[H^+]$: Calc 396.2064; Found 396.2065. 1H NMR (300 MHz, CD_3OD) δ 8.32 (s, 1H, **1 or 2**), 8.21 (s, 1H, **1 or 2**), 6.00 (d, $J = 4.8$ Hz, 1H, **6**), 4.81 – 4.75 (m, 1H, **8**), 4.37 – 4.30 (m, 1H, **9**), 4.23 – 4.16 (m, 1H, **7**), 2.96 – 2.89 (m, 1H, **13a**), 2.60 – 2.48 (m, 1H, **13b**), 1.63 – 1.15 (m, 14H, **10**), 0.88 (t, $J = 6.1$ Hz, 3H, **15**). ^{13}C NMR (75 MHz, CD_3OD) δ 155.9 (s), 152.5 (s), 149.3 (s), 140.0 (s), 119.1 (s), 88.7 (s), 84.5 (s), 73.5 (s), 72.6 (s), 33.8 (s), 32.5 (s), 31.6 (s), 29.6 (s), 28.9 (s), 28.8 (s), 28.5 (s), 22.3 (s), 13.0 (s).

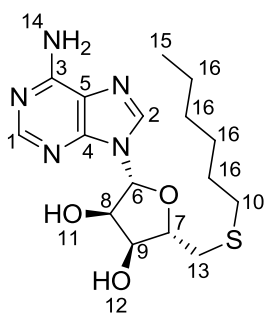
23 9-((3aR,4R,6S,6aS)-6-((Hexylthio)methyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine



To a solution, cooled to -15 °C, of **20** (100 mg, 0.274 mmol) in freshly distilled MeOH was added a solution of NaOMe in MeOH 25% m/v (64 μ l, 0.301 mmol). The reaction was allowed to proceed for 45 minutes, after which 1-bromohexane (90.3 mg, 0.547 mmol) was added. The reaction was allowed to proceed for 24 hours, after it was quenched by the addition of water (5 mL). The MeOH was removed *in vacuo* and the remainder was extracted with DCM (3 x 150 mL). The recombined organic layers were washed with brine and dried with $MgSO_4$. The solvent was removed *in vacuo* and the crude was lyophilised overnight.

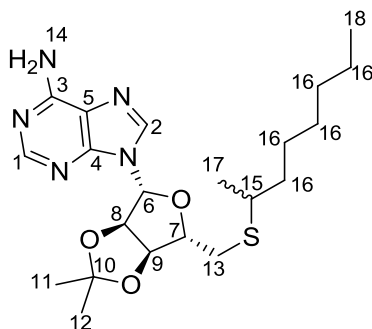
Flash chromatography (DCM to 4:96 MeOH:DCM) afforded the pure oil (55 mg, 63%). $R_f = 0.56$ (1:9 MeOH:DCM). MS Elemental: $C_{19}H_{30}N_5O_3S$ [H^+]: Calc 408.2064; Found 408.2069. 1H NMR (300 MHz, $CDCl_3$) δ 8.34 (s, 1H, **1 or 2**), 7.92 (s, 1H, **1 or 2**), 6.08 (d, $J = 2.2$ Hz, 1H, **6**), 5.84 (s, 2H, **14**), 5.51 (dd, $J = 6.4, 2.2$ Hz, 1H, **8**), 5.05 (dd, $J = 6.4, 3.1$ Hz, 1H, **9**), 4.39 (ddd, $J = 7.5, 6.3, 3.1$ Hz, 1H, **7**), 2.83 (dd, $J = 13.6, 7.5$ Hz, 1H, **13a**), 2.75 (dd, $J = 13.6, 6.3$ Hz, 1H, **13b**), 2.52 – 2.45 (m, 2H, **15**), 1.61 (s, 3H, **11 or 12**), 1.56 – 1.43 (m, 2H, **16**), 1.39 (s, 3H, **11 or 12**), 1.38 – 1.13 (m, 6H, **16**), 0.92 – 0.80 (m, 3H, **17**). ^{13}C NMR (75 MHz, $CDCl_3$) δ 155.7 (s), 153.3 (s), 149.4 (s), 140.1 (s), 120.4 (s), 114.57 (s), 91.1 (s), 87.0 (s), 84.2 (s), 84.0 (s), 34.5 (s), 32.8 (s), 31.5 (s), 29.7 (s), 28.6 (s), 27.2 (s), 25.5 (s), 22.7 (s), 14.2 (s).

24 (2R,3R,4S,5S)-2-(6-Amino-9H-purin-9-yl)-5-((hexylthio)methyl)tetrahydrofuran-3,4-diol



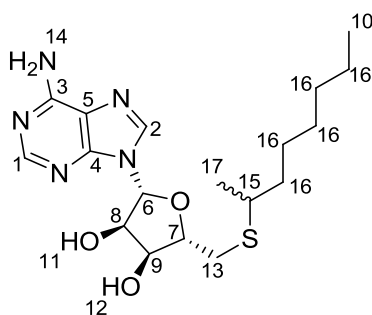
To a solution of **23** (55 mg, 0.136 mmol) in water (7 ml) was added formic acid (2 ml). The reaction was left to stir for 4 days. Volatiles were removed *in vacuo* to yield pure white powder (50.02 mg, 100%). $R_f = 0.63$ (1:9 MeOH:DCM). MS Elemental: $C_{16}H_{26}N_5O_3S$ [H^+]: Calc 368.1751; Found 368.1756. 1H NMR (300 MHz, CD_3OD) δ 8.33 (s, 1H, **1 or 2**), 8.22 (s, 1H, **1 or 2**), 6.02 (d, $J = 4.9$ Hz, 1H, **6**), 4.82 – 4.77 (m, 1H, **8**), 4.39 – 4.33 (m, 1H, **9**), 4.26 – 4.18 (m, 1H, **7**), 2.98 (dd, $J = 14.2, 5.4$ Hz, 1H, **13a**), 2.90 (dd, $J = 14.2, 6.0$ Hz, 1H, **13b**), 2.58 – 2.49 (m, 2H, **10**), 1.60 – 1.46 (m, 2H, **16**), 1.40 – 1.13 (m, 6H, **16**), 0.91 – 0.83 (m, 3H, **15**). ^{13}C NMR (75 MHz, CD_3OD) δ 157.3 (s), 153.8 (s), 150.7 (s), 141.4 (s), 120.5 (s), 90.1 (s), 85.9 (s), 74.9 (s), 74.0 (s), 35.2 (s), 33.9 (s), 32.5 (s), 30.7 (s), 29.5 (s), 23.6 (s), 14.4 (s).

25 9-((3a*R*,4*R*,6*S*,6a*S*)-2,2-Dimethyl-6-((octan-2-ylthio)methyl)tetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine



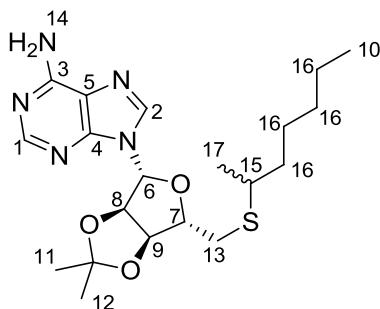
To a solution, cooled to $-15\text{ }^{\circ}\text{C}$, of **20** (50 mg, 0.137 mmol) in freshly distilled MeOH (5 mL) was added NaOMe in MeOH 25% m/v (30 μl , 0.137 mmol). After 45 minutes, 2-bromooctane (47.7 mg, 0.247 mmol) was added and allowed to stir for 24 hours. The reaction was quenched with water and the solvent removed. After the crude was lyophilised overnight, flash chromatography (1:99 MeOH:DCM to 4:96 MeOH:DCM) afforded an oil (25 mg, 41%). $R_f = 0.56$ (1:9 MeOH:DCM). MS Elemental: $\text{C}_{21}\text{H}_{34}\text{N}_5\text{O}_3\text{S}$ [H^+]: Calc 436.2377; Found 436.2382. ^1H NMR (300 MHz, CDCl_3) δ 8.35 (s, 1H, **1 or 2**), 7.93 (s, 1H, **1 or 2**), 6.07 (d, $J = 2.1$ Hz, 1H, **6**), 5.86 – 5.72 (m, 2H, **14**), 5.53 – 5.48 (m, 1H, **8**), 5.09 – 5.03 (m, 1H, **9**), 4.42 – 4.34 (m, 1H, **7**), 2.91 – 2.81 (m, 1H, **13a**), 2.81 – 2.67 (m, 1H, **13b**), 1.66 – 1.14 (m, 20H, **11, 12, 15, 16 and 17**), 0.89 – 0.81 (m, 3H, **18**).

26 (2*R*,3*R*,4*S*,5*S*)-2-(6-Amino-9*H*-purin-9-yl)-5-((octan-2-ylthio)methyl)tetrahydrofuran-3,4-diol



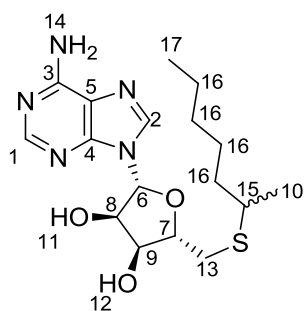
To a solution of **25** (25 mg, 0.057 mmol) in water (7 mL) was added formic acid (2 mL). The reaction was left to stir for 4 days. Volatiles were removed *in vacuo* to yield pure white powder (23.43 mg, 100%). $R_f = 0.63$ (1:9 MeOH:DCM). MS [H^+]: Calc 396.2064; Found 396.2067. ^1H NMR (300 MHz, CD_3OD) δ 8.34 – 8.30 (m, 1H, **1 or 2**), 8.21 (s, 1H, **1 or 2**), 6.00 (d, $J = 4.9$ Hz, 1H, **6**), 4.83 – 4.76 (m, 1H, **8**), 4.38 – 4.30 (m, 1H, **9**), 4.24 – 4.16 (m, 1H, **7**), 3.04 – 2.95 (m, 1H, **13a**), 2.95 – 2.85 (m, 1H, **13b**), 2.85 – 2.75 (m, 1H, **15**), 1.61 – 1.12 (m, 13H, **16 and 17**), 0.87 (t, $J = 6.8$ Hz, 3H, **10**).

27 9-((3a*R*,4*R*,6*S*,6a*S*)-6-((Heptan-2-ylthio)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine

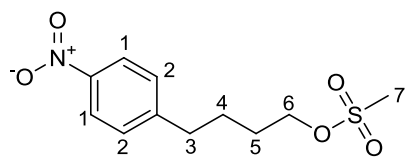


To a -15°C solution of **20** (40 mg, 0.109 mmol) in freshly distilled MeOH, was added dropwise NaOMe in MeOH 25% m/v (23.6 μl , 0.109 mmol) and allowed to stir. After 45 minutes, 2-bromoheptane (98.0 mg, 0.547 mmol) was added and allowed to stir for 24 hours. The reaction was quenched with water and the solvent was removed *in vacuo*. After the crude was lyophilised overnight, flash chromatography (DCM to 4:96 MeOH:DCM) yielded the product as a pale yellow oil (38 mg, 82%). $R_f = 0.56$ (1:9 MeOH:DCM). MS Elemental: $\text{C}_{20}\text{H}_{32}\text{N}_5\text{O}_3\text{S}$ [H^+]: Calc 422.2220; Found 422.2226. ^1H NMR (300 MHz, CDCl_3) δ 8.34 (s, 1H, **1 or 2**), 7.93 (s, 1H, **1 or 2**), 6.07 (d, $J = 2.2$ Hz, 1H, **6**), 5.93 (s, 2H, **14**), 5.54 – 5.46 (m, 1H, **8**), 5.10 – 5.02 (m, 1H, **9**), 4.43 – 4.33 (m, 1H, **7**), 2.90 – 2.66 (m, 3H, **13a**, **13b** and **15**), 1.64 – 1.17 (m, 17H, **11**, **12**, **16** and **10 or 17**), 0.90 – 0.82 (m, 3H, **10 or 17**).

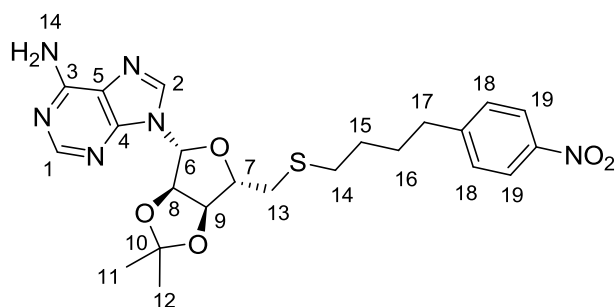
28 (2*R*,3*R*,4*S*,5*S*)-2-(6-Amino-9*H*-purin-9-yl)-5-((heptan-2-ylthio)methyl) tetrahydrofuran-3,4-diol



To a solution of **27** (25.9 mg, 0.061) in water (7 ml) was added formic acid (2 ml). The reaction was left to stir for 4 days. Volatiles were removed *in vacuo* to yield pure white powder (23.43 mg, 99%). $R_f = 0.63$ (1:9 MeOH:DCM). MS Elemental: $\text{C}_{17}\text{H}_{28}\text{N}_5\text{O}_3\text{S}$ [H^+]: Calc 382.1907; Found 382.1913. ^1H NMR (300 MHz, CD_3OD) δ 8.33 – 8.28 (m, 1H, **1 or 2**), 8.20 (s, 1H, **1 or 2**), 6.00 (d, $J = 4.8$ Hz, 1H, **6**), 4.82 – 4.76 (m, 1H, **8**), 4.39 – 4.30 (m, 1H, **9**), 4.24 – 4.14 (m, 1H, **7**), 3.03 – 2.94 (m, 1H, **13a**), 2.94 – 2.84 (m, 1H, **13b**), 2.84 – 2.72 (m, 1H, **15**), 1.60 – 1.11 (m, 11H, **16** and **10 or 17**), 0.89 – 0.81 (m, 3H, **10 or 17**).

33 4-(4-Nitrophenyl)butyl methanesulfonate

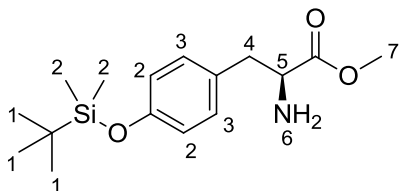
To an ice cold solution of 4-(4-nitrophenyl)butan-1-ol **34** (500 mg, 2.56 mmol) and TEA (499 μl , 3.59 mmol) in DCM (15 ml) was added dropwise mesylchloride (237 μl , 3.07 mmol). This was allowed to stir for 3 hours. The suspension was filtered and the filtrate concentrated. The concentrate was purified by flash chromatography to yield the pure oil (635 mg, 90%). ^1H NMR (300 MHz, CDCl_3) δ 8.19 – 8.12 (m, 2H, **1 or 2**), 7.38 – 7.30 (m, 2H, **1 or 2**), 4.29 – 4.22 (m, 2H, **3 or 6**), 3.01 (s, 3H, **7**), 2.82 – 2.73 (m, 2H, **3 or 6**), 1.88 – 1.71 (m, 4H, **4 and 5**).

32 9-((3aR,4R,6S,6aS)-2,2-Dimethyl-6-(((4-(4-nitrophenyl)butyl)thio)methyl) tetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine

To a solution, cooled to $-15\text{ }^\circ\text{C}$, of **20** (300 mg, 0.821 mmol) in freshly distilled MeOH (15 ml) was added dropwise addition of a solution of NaOMe in MeOH 25% m/v (203 μl , 0.944 mmol) and allowed to stir for 45 minutes. This was followed by the addition of **33** (258 mg, 0.944 mmol). The reaction was then allowed to stir for 24 hours, after which the reaction was quenched by the addition of water. The MeOH was removed *in vacuo* and the water layer was then extracted with DCM (4x 100 ml). The recombined organic layers were washed with brine, dried with MgSO_4 and filtered. The crude was purified by flash chromatography (DCM to 2:98 MeOH:DCM to 4:94 MeOH:DCM) to yield product (65 mg, 15.8%). MS Elemental: $\text{C}_{23}\text{H}_{29}\text{N}_6\text{O}_5\text{S}$ [H^+]: Calc 501.1914; Found 501.1920. ^1H NMR (300 MHz, CDCl_3) δ 8.31 (s, 1H, **1 or 2**), 8.13 – 8.06 (m, 2H, **18 or 19**), 7.91 (s, 1H, **1 or 2**), 7.30 – 7.24 (m, 2H, **18 or 19**), 6.10 (s, 2H, **14**), 6.07 (d, $J = 2.2$ Hz, 1H, **6**), 5.51 (dd, $J = 6.4, 2.2$ Hz, 1H, **8**), 5.05 (dd, $J = 6.4, 3.2$ Hz, 1H, **9**), 4.36 (td, $J = 6.9, 3.2$ Hz, 1H, **7**), 2.82 (dd, $J = 13.7, 7.2$ Hz, 1H, **13a**), 2.74 (dd, $J = 13.7, 6.6$ Hz, 1H, **13b**), 2.64 (m, 2H, **14, 15 or 16**), 2.50 (t, $J = 7.1$ Hz, 2H, **17**), 1.72 – 1.44 (m, 7H, **14, 15 or 16 and 11 or 12**), 1.38 (s, 3H, **11 or 12**). ^{13}C NMR (75 MHz, CDCl_3) δ 155.8 (s), 153.2 (s), 150.0 (s), 149.3 (s), 146.4 (s), 140.2 (s), 129.2 (s), 123.7 (s),

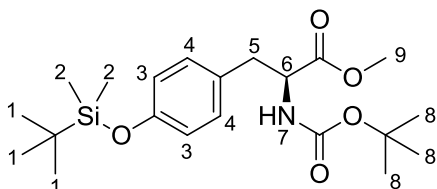
120.4 (s), 114.6 (s), 91.0 (s), 87.1 (s), 84.1 (s), 84.0 (s), 35.4 (s), 34.5 (s), 32.4 (s), 29.9 (s), 29.0 (s), 27.2 (s), 25.4 (s).

39 (S)-Methyl 2-amino-3-(4-((tert-butyl)dimethylsilyloxy)phenyl)propanoate



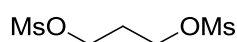
To a stirred solution of L-tyrosine methyl ester **37** (500 mg, 2.56 mmol) in THF (50 mL) was added imidazole (261 mg, 3.84 mmol) and (463 mg, 3.07 mmol) TBDMSCl. After 16 hours, the solvent was removed *in vacuo* and was partitioned between saturated NaHCO₃ and DCM (3 x 100 mL). After washing with brine, drying with MgSO₄ and filtration, the crude was purified by flash chromatography (1:1 EtOAc:Hexanes) to yield a colourless oil (502 mg, 63%). R_f = 0.35 (1:1 EtOAc:Hexanes). ¹H NMR (400 MHz, CDCl₃) δ 7.09 – 6.96 (m, 2H, **2 or 3**), 6.81 – 6.71 (m, 2H, **2 or 3**), 3.70 (s, 3H, **7**), 3.69 – 3.66 (m, 1H, **5**), 3.07 – 2.73 (m, 2H, **4**), 1.49 (s, 2H, **6**), 0.97 (s, 9H, **1**), 0.18 (s, 6H, **2**).

40 (S)-Methyl 2-((tert-butoxycarbonyl)amino)-3-(4-((tert-butyl)dimethylsilyloxy)phenyl)propanoate



To a solution of **39** (145 mg, 0.469 mmol) in MeCN (40 mL), was added Boc₂O (102 mg, 0.469 mmol) and a catalytic amount of DMAP. The reaction was then left to proceed for 16 hours, after which the solvent was removed *in vacuo*. Purification by column chromatography (1:9 EtOAc:Hexanes) afforded an oil (15 mg, 7%). R_f = 0.61 (2:8 EtOAc:Hexanes). ¹H NMR (300 MHz, CDCl₃) δ 7.01 – 6.93 (m, 2H, **3 or 4**), 6.81 – 6.71 (m, 2H, **3 or 4**), 5.00 – 4.88 (m, 1H, **6 or 7**), 4.59 – 4.46 (m, 1H, **6 or 7**), 3.69 (s, 3H, **9**), 3.05 – 2.93 (m, 2H, **5**), 1.42 (s, 9H, **1 or 8**), 0.97 (s, 9H, **1 or 8**), 0.18 (s, 6H, **2**).

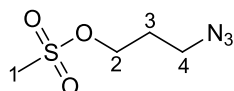
44 Propane-1,3-diyl dimethanesulfonate



To an ice cold solution of propane-1,3-diol **43** (500 mg, 6.57 mmol) in THF (20 mL) was added (2289 μL TEA, 16.43 mmol). This was followed by the dropwise addition of MsCl

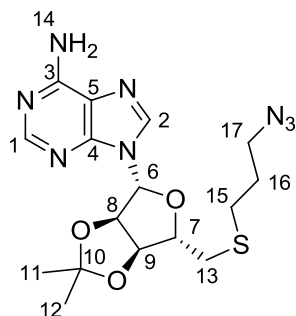
(1220 μl , 15.77 mmol) and was allowed to stir for 4 hours. The white precipitate was removed by filtration and the yellow filtrate was concentrated *in vacuo*. The crude was used without further purification.

45 3-Azidopropyl methanesulfonate



To a stirred solution of **44** (1429 mg, 6.152 mmol) in MeCN was added NaN_3 (400 mg, 6.152 mmol) and set to reflux for 16 hours. The solvent was removed *in vacuo* and the crude was partitioned against water and EtOAc (3x 100 mL). Afterwards, the recombined organic layers were washed with brine, dried with MgSO_4 and filtered. The crude was purified by flash chromatography (1:4 EtOAc:Hexanes) to yield a colourless oil over two steps (494 mg, 42%). $R_f = 0.42$ (1:4 EtOAc:Hexanes). $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 4.38 – 4.27 (t, 2H, **2 or 4**), 3.53 – 3.41 (m, 2H, **2 or 4**), 3.04 (s, 3H, **1**), 2.07 – 1.94 (m, 2H, **3**).

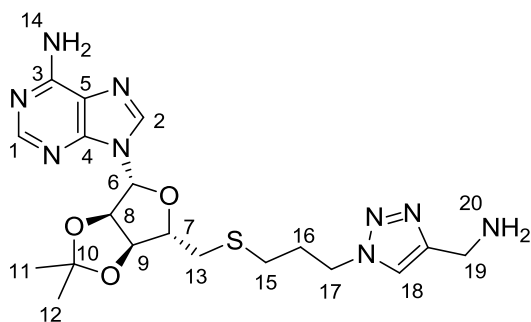
46 9-((3*R*,4*R*,6*S*,6*S*)-6-(((3-Azidopropyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine



To a solution, cooled to $-15\text{ }^\circ\text{C}$, of **20** (300 mg, 0.821 mmol) in freshly distilled MeOH (10 mL) was added NaOMe in MeOH 25% m/v (185 μl , 0.858 mmol) and allowed to proceed for 1 hour. This was followed by the addition of **45** (294 mg, 1.642 mmol) and allowed to stir for a 24 hours. The reaction was quenched with water. After the MeOH was removed *in vacuo*, the solution was extracted with chloroform (5 x 100 mL), washed with brine and dried with MgSO_4 . The solution was concentrated *in vacuo* and purified by flash chromatography (DCM to 4:96 MeOH:DCM) to yield pure product (145 mg, 65%). $R_f = 0.63$ (1:9 MeOH:DCM). MS Elemental: $\text{C}_{16}\text{H}_{21}\text{N}_8\text{O}_3\text{S}$ [H^+]: Calc 407.1608; Found 407.1614. $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.33 (s, 1H, **1 or 2**), 7.91 (s, 1H, **1 or 2**), 6.07 (d, $J = 2.1$ Hz, 1H, **6**), 6.03 (s, 2H, **14**), 5.51 (dd, $J = 6.4, 2.1$ Hz, 1H, **8**), 5.06 (dd, $J = 6.4, 3.2$ Hz, 1H, **9**), 4.37 (td, $J = 6.9, 3.2$ Hz, 1H, **7**), 3.33 (t, $J = 6.6$ Hz, 2H, **17**), 2.85 (dd, $J = 13.7, 7.3$ Hz, 1H, **13a**), 2.78 (dd, $J = 13.7, 6.5$ Hz, 1H, **13b**), 2.59 – 2.53 (m, $J = 7.1$ Hz, 2H, **15**), 1.75 (m, 2H, **16**), 1.60 (s, 3H, **11 or 12**), 1.39

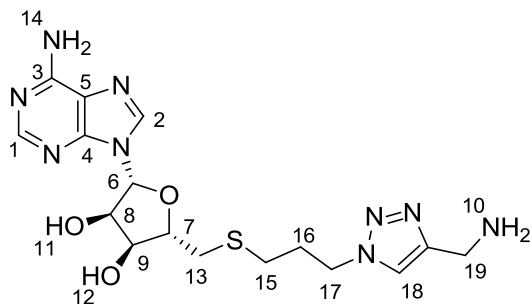
(s, 3H, **11 or 12**). ^{13}C NMR (101 MHz, CDCl_3) δ 155.8 (s), 153.4 (s), 149.4 (s), 140.3 (s), 120.6 (s), 114.7 (s), 91.1 (s), 87.1 (s), 84.2 (s), 84.0 (s), 50.1 (s), 34.6 (s), 29.8 (s), 28.9 (s), 27.3 (s), 25.6 (s).

47 9-((3a*R*,4*R*,6*S*,6a*S*)-6-(((3-(4-(Aminomethyl)-1*H*-1,2,3-triazol-1-yl)propyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine



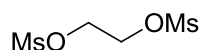
To a solution of **46** (148 mg, 0.364 mmol) in MeCN (5 mL) was added TEA (110 μl , 1.09 mmol) and propargylamine **42** (22.2 mg, 0.437 mmol). Finally, CuI (6.4 mg, 0.034 mmol) was added and the reaction proceeded for 1 hour. An additional portion of propargyl amine (9 mg, 0.161 mmol) was added and the reaction proceeded for an additional hour. Saturated ammonium chloride was added and the organics were extracted with chloroform (6 x 50 mL). The solvent was removed *in vacuo* and the distillate was purified using column chromatography (10:2:88 MeOH:TEA:DCM) to yield pure product (140 mg, 83%). R_f = 0.71 (4:96 MeOH:DCM). MS Elemental: $\text{C}_{19}\text{H}_{28}\text{N}_9\text{O}_3\text{S}$. $[\text{H}^+]$: calc: 462.2030; found: 462.2044. ^1H NMR (400 MHz, CDCl_3) δ 8.27 (s, 1H, **1 or 2**), 7.90 (s, 1H, **1 or 2**), 7.41 (s, 1H, **18**), 6.38 (s, 2H, **14**), 6.05 (d, J = 2.0 Hz, 1H, **6**), 5.50 (dd, J = 6.4, 2.0 Hz, 1H, **8**), 5.04 (dd, J = 6.4, 3.1 Hz, 1H, **9**), 4.36 – 4.29 (m, 3H, **7 and 17**), 3.95 (s, 2H, **19**), 3.05 (s, 3H, **20**), 2.81 (dd, J = 13.7, 6.9 Hz, 1H, **13a**), 2.75 (dd, J = 13.7, 6.8 Hz, 1H, **13b**), 2.43 (t, J = 7.0 Hz, 2H, **15**), 2.06 – 1.97 (m, 2H, **16**), 1.57 (s, 3H, **11 or 12**), 1.36 (s, 3H, **11 or 12**). ^{13}C NMR (101 MHz, CDCl_3) δ 155.9 (s), 153.2 (s), 149.1 (s), 140.1 (s), 121.2 (s), 114.5 (s), 90.9 (s), 87.2 (s), 84.0 (s), 83.9 (s), 48.5 (s), 46.0 (s), 37.5 (s), 34.4 (s), 29.7 (s), 29.3 (s), 27.1 (s), 25.4 (s), 10.7 (s).

48 (2R,3R,4S,5S)-2-(6-Amino-9H-purin-9-yl)-5-(((3-(4-(aminomethyl)-1H-1,2,3-triazol-1-yl)propyl)thio)methyl)tetrahydrofuran-3,4-diol



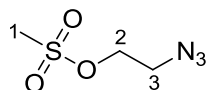
To a solution of formic acid and water was added **47** (16 mg, 0.035 mmol) and was allowed to stir for 3 days. The volatiles were then removed *in vacuo* and the remaining solvent was removed by lyophilisation to yield a white powder (14 mg, 99%). $R_f = 0.13$ (1:4 MeOH:DCM). MS Elemental: $C_{16}H_{24}N_9O_3$ [H^+]: Calc 422.1717; Found 422.1723. 1H NMR (300 MHz, CD_3OD) δ 8.15 (s, 1H, **1, 2 or 18**), 8.05 (s, 1H, **1, 2 or 18**), 7.87 (s, 1H, **1, 2 or 18**), 5.85 (d, $J = 4.8$ Hz, 1H, **6**), 4.70 – 4.64 (m, 1H, **8**), 4.32 (t, $J = 6.8$ Hz, 2H, **17**), 4.23 – 4.16 (m, 1H, **9**), 4.11 (s, 2H, **19**), 4.06 – 3.99 (m, 1H, **7**), 3.20 – 3.14 (m, 1H, **13a**), 2.85 – 2.78 (m, 1H, **13b**), 2.40 (t, $J = 6.9$ Hz, 2H, **15**), 2.05 – 1.91 (m, 2H, **16**). ^{13}C NMR (75 MHz, CD_3OD) δ 165.6 (s), 157.1 (s), 153.5 (s), 150.6 (s), 141.6 (s), 141.2 (s), 125.5 (s), 120.5 (s), 90.2 (s), 85.7 (s), 74.8 (s), 74.1 (s), 35.4 (s), 35.1 (s), 30.9 (s), 30.2 (s).

50 Ethane-1,2-diyl dimethanesulfonate



To an ice cold solution of ethane-1,2-diol **49** (500 mg, 8.06 mmol) in THF (15 ml) was added (1933 μ l TEA, 20.14 mmol). This was followed by the dropwise addition of MsCl (1030 μ l, 19.33 mmol) and was allowed to stir for 16 hours. The white precipitate was removed by filtration and the yellow filtrate was concentrated *in vacuo*. The crude was used without further purification.

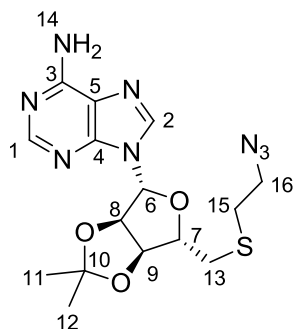
51 2-Azidoethyl methanesulfonate



To a stirred solution of **50** (1129 mg, 8.06 mmol) in MeCN was added NaN_3 (524 mg, 8.06 mmol) and set to reflux for 16 hours. The solvent was removed *in vacuo* and the crude was partitioned against water and EtOAc (3x 100 ml). Afterwards, the recombined organic layers were washed with brine, dried with $MgSO_4$ and filtered. The crude was purified by flash

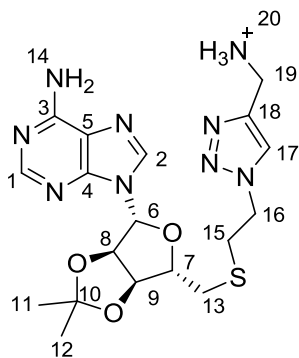
chromatography (2:8 EtOAc:Hexanes) to yield a colourless oil over two steps (628 mg, 47%). $R_f = 0.42$ (2:8 EtOAc:Hexanes). $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 4.39 – 4.30 (m, 2H, **2 or 3**), 3.63 – 3.55 (m, 2H, **2 or 3**), 3.11 – 3.05 (m, 3H, **1**).

52 9-((3aR,4R,6S,6aS)-6-(((2-Azidoethyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine

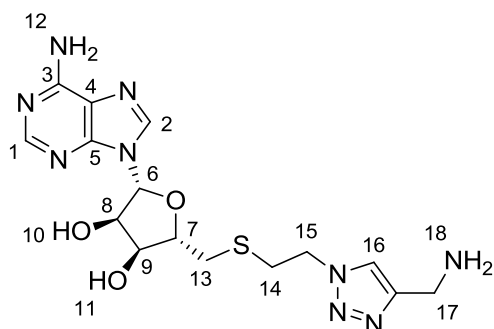


To a solution, cooled to $-15\text{ }^\circ\text{C}$, of **20** (300 mg, 0.821 mmol) in freshly distilled MeOH (10 mL) was added NaOMe in MeOH 25% m/v (185 μl , 0.858 mmol) and allowed to proceed for 1 hour. This was followed by the addition of **51** (271 mg, 1.641 mmol) and allowed to stir for a 24 hours. The reaction was quenched with water. After the MeOH was removed *in vacuo*, the solution was extracted with CHCl_3 (5x 100 mL), washed with brine, dried with MgSO_4 and filtered. The solution was concentrated *in vacuo* and purified by flash chromatography (DCM to 1:24 MeOH/DCM) to yield pure product (110 mg, 34%). $R_f = 0.63$ (1:9 MeOH:DCM). MS Elemental: $\text{C}_{15}\text{H}_{21}\text{N}_8\text{O}_3\text{S}$ [H^+]: Calc 393.1451; Found 393.1457. $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 8.32 (s, 1H, **1 or 2**), 7.90 (s, 1H, **1 or 2**), 6.32 (s, 2H, **14**), 6.07 (d, $J = 2.2$ Hz, 1H, **6**), 5.50 (dd, $J = 6.4, 2.2$ Hz, 1H, **8**), 5.07 (dd, $J = 6.4, 3.3$ Hz, 1H, **9**), 4.41 – 4.33 (m, 1H, **7**), 3.35 (t, $J = 6.8$ Hz, 2H, **15 or 16**), 2.95 – 2.79 (m, 2H, **13**), 2.66 (t, $J = 6.8$ Hz, 2H, **15 or 16**), 1.59 (s, 3H, **11 or 12**), 1.38 (s, 3H, **11 or 12**). $^{13}\text{C NMR}$ (75 MHz, CDCl_3) δ 156.0 (s), 153.2 (s), 149.2 (s), 140.1 (s), 120.4 (s), 114.7 (s), 90.9 (s), 87.2 (s), 84.1 (s), 83.8 (s), 51.1 (s), 34.6 (s), 31.8 (s), 27.2 (s), 25.4 (s).

53 (1-(2-(((3a*S*,4*S*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro [3,4-*d*][1,3]dioxol-4-yl)methyl)thio)ethyl)-1*H*-1,2,3-triazol-4-yl)methanaminium



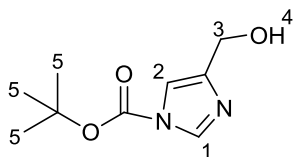
To a solution of **52** (110 mg, 0.280 mmol) in MeCN (5 mL) was added TEA (119 μ l, 0.841 mmol) and propargylamine **42** (18.5 mg, 0.336 mmol). Finally, CuI (5.3 mg, 0.028 mmol) was added and the reaction stirred for 1 hour. An additional portion of propargyl amine (9mg, 0.161 mmol) was added and the reaction proceeded for an additional hour. Saturated ammonium chloride was added and the organics were extracted with chloroform (6x 50 mL). The solvent was removed *in vacuo* and the distillate was purified using column chromatography (10:2:88 MeOH:TEA:DCM), a white product was obtained (43 mg, 34%). R_f = 0.23 (10:1:89 MeOH:TEA:DCM). MS [H^+]: calc: 434.1717; found: 434.1723. 1H NMR (300 MHz, $CDCl_3$) δ 8.31 (s, 1H, **1 or 2**), 7.91 (s, 1H, **1 or 2**), 7.35 (s, $J = 7.5$ Hz, 1H, **1 or 17**), 6.16 (s, 2H, **14**), 6.05 (d, $J = 2.1$ Hz, 1H, **6**), 5.49 (dd, $J = 6.4, 2.1$ Hz, 1H, **8**), 5.04 (dd, $J = 6.4, 3.4$ Hz, 1H, **9**), 4.48 – 4.25 (m, 3H, **7 and -CH₂-**), 3.95 (s, 2H, **19**), 3.17 – 3.00 (m, 3H, **20**), 2.96 – 2.86 (m, 2H, **13**), 2.83 – 2.77 (m, 2H, **-CH₂-**), 1.59 (s, 3H, **11 or 12**), 1.37 (s, 3H, **11 or 12**). ^{13}C NMR (75 MHz, CD_3OD) δ 155.9 (s), 153.3 (s), 149.2 (s), 140.2 (s), 121.3 (s), 120.4 (s), 114.7 (s), 90.9 (s), 87.4 (s), 84.0 (s), 83.8 (s), 53.2 (s), 50.6 (s), 49.9 (s), 34.6 (s), 32.7 (s), 27.2 (s), 25.4 (s).

54 (2R,3R,4S,5S)-2-(6-Amino-9H-purin-9-yl)-5-(((2-(4-(aminomethyl)-1H-1,2,3-triazol-1-yl)ethyl)thio)methyl)tetrahydrofuran-3,4-diol

To a solution of formic acid and water was added **53** (10.6 mg, 0.024 mmol) to stir for 3 days. The volatiles were removed *in vacuo* and the remaining solvent was removed by lyophilisation to yield a white powder (9.6 mg, 99%). $R_f = 0.11$ (1:4 MeOH:DCM). MS Elemental: $C_{15}H_{22}N_9O_3S$. $[H^+]$: Calc 408.1561; Found 408.1566. 1H NMR (400 MHz, CD_3OD) δ 8.24 – 8.21 (m, 1H, **1**, **2** or **16**), 8.15 (d, $J = 2.0$ Hz, 1H, **1**, **2** or **16**), 7.95 (s, 1H, **1**, **2** or **16**), 5.95 (d, $J = 4.7$ Hz, 1H, **6**), 4.75 (t, $J = 5.1$ Hz, 1H, **8**), 4.55 – 4.50 (m, 2H, **-CH₂-**), 4.28 (t, $J = 5.2$ Hz, 1H, **9**), 4.19 (s, 2H, **17**), 4.15 – 4.09 (m, 1H, **7**), 3.09 – 3.02 (m, 1H, **13a**), 3.02 – 2.95 (m, 1H, **13b**), 2.94 – 2.90 (m, 2H, **-CH₂-**). ^{13}C NMR (75 MHz, CD_3OD) δ 165.6 (s), 153.8 (s), 141.6 (s), 125.7 (s), 90.3 (s), 86.0 (s), 74.7 (s), 74.0 (s), 51.0 (s), 47.8 (s), 35.4 (s), 35.2 (s), 33.7 (s), 9.2 (s), 7.6 (s).

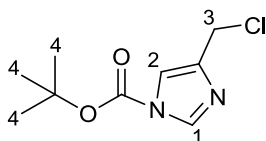
Synthetic Procedures (Chapter 7)

58 *Tert*-butyl 4-(hydroxymethyl)-1*H*-imidazole-1-carboxylate

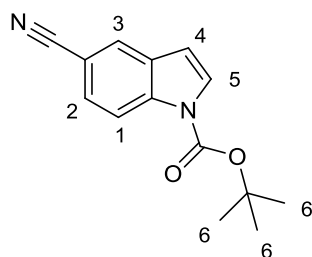


To a solution of (1*H*-imidazol-4-yl)methanol **59** (500 mg, 5.10 mmol) in DMF (20 mL) was added di-*tert*-butyl dicarbonate (1334 mg, 6.11 mmol). To this was added a single crystal of DMAP and left to stir for 1 hour at RT. The solvent removed *in vacuo* by toluene azeotrope. The remainder was partitioned against brine and EtOAc (3 x 50 mL). After removal of the solvent *in vacuo*, the crude was purified by column chromatography (35:65 EtOAc:Hexanes to 70:30 EtOAc:Hexanes to EtOAc). $R_f = 0.35$ (35:65 EtOAc:Hexanes). $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 8.06 – 7.98 (m, 1H, **1 or 2**), 7.26 (s, 1H, **2 or 1**), 4.55 (s, 2H, **3**), 2.37 (s, 1H, **4**), 1.65 – 1.57 (m, 9H, **5**).

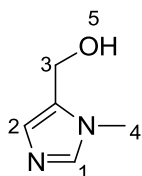
57 *Tert*-butyl 4-(chloromethyl)-1*H*-imidazole-1-carboxylate



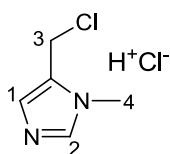
To an ice cold solution of **58** (300 mg, 1.51 mmol) in DCM (7 mL) and 1 drop DMF, was added dropwise a solution of SOCl_2 (0.200 mL, 2.72 mmol) in DCM (0.35 mL) over 3 minutes. The reaction was left to stir for 20 minutes after which the solvent was removed *in vacuo*. The crude was dissolved in EtOAc, and rinsed with ice cold saturated NaHCO_3 . The aqueous phase was extracted with (2x 50 mL) EtOAc. The recombined organic layers were washed with brine and dried using NaSO_4 . Flash chromatography (2:8 EtOAc:Hexanes) afforded a colourless oil (270 mg, 83%). $R_f = 0.32$ (2:8 EtOAc:Hexanes). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.05 – 8.02 (m, 1H, **1 or 2**), 7.37 (s, 1H, **2 or 1**), 4.56 – 4.49 (m, 2H, **3**), 1.61 – 1.59 (m, 9H, **4**).

61 Tert-butyl 5-cyano-1H-indene-1-carboxylate

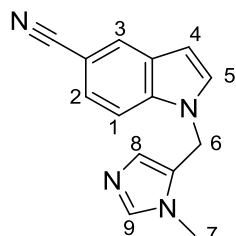
To an ice cold solution of 1H-indole-5-carbonitrile **56** (63.0 mg, 0.443 mmol) in DMF (2 mL) was added NaH in oil dispersion (15.6 mg, 0.650 mmol). After evolution of hydrogen ceased (5 minutes), **57** (120 mg, 0.553 mmol) dissolved in DMF (3 mL) was added. The reaction was allowed to proceed for 10 minutes. The reaction was quenched with the addition of water and extracted with EtOAc (3 x 50 mL). Afterwards the recombined organic layers were washed with brine, dried with MgSO₄ and filtered. Flash chromatography (1:9 EtOAc:Hexanes) afforded the pure product (52.9 mg, 49.2%). R_f = 0.28 (1:9 EtOAc:Hexanes). ¹H NMR (300 MHz, CDCl₃) δ 8.25 (d, *J* = 8.6 Hz, 1H, **1**), 7.89 (d, *J* = 1.6 Hz, 1H, **3**), 7.70 (d, *J* = 3.8 Hz, 1H, **4 or 5**), 7.55 (dd, *J* = 8.6, 1.6 Hz, 1H, **2**), 6.63 (d, *J* = 3.8 Hz, 1H, **4 or 5**), 1.68 (s, 9H, **6**).

64 (1-Methyl-1H-imidazol-5-yl)methanol

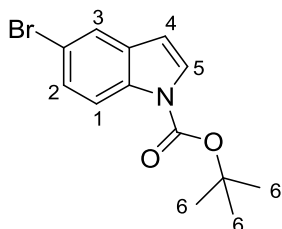
To an ice cold solution of 1-methyl-1H-imidazole-5-carbaldehyde **63** (275 mg, 2.50 mmol) in 10 mL THF was added LiAlH₄ (220 mg, 5.80 mmol) and allowed to stir for 20 minutes. The reaction was quenched by the careful addition of NaSO₄•10H₂O and allowed to stir for an additional hour. The precipitate was removed via filtration and the solvent was removed *in vacuo* to yield product (180 mg, 88.6%). ¹H NMR (300 MHz, CDCl₃) δ 7.36 (s, 1H, **1 or 2**), 6.84 (s, 1H, **2 or 1**), 4.60 (s, 2H, **3**), 3.69 (s, 3H, **4**), 3.59 – 2.10 (m, 1H, **5**).

65 5-(Chloromethyl)-1-methyl-1H-imidazole hydrochloride

To a stirred solution of **64** (171 mg, 1.52 mmol) in SOCl_2 (5 mL) was added 1 drop of DMF. After 30 minutes, the solvent was removed *in vacuo*. The remainder was dissolved in ethanol and triturated with diethyl ether to yield an off white solid (179 mg, 90%). ^1H NMR (300 MHz, DMSO) δ 9.21 – 9.18 (m, 1H, **1 or 2**), 7.80 – 7.78 (m, 1H, **2 or 1**), 5.04 – 5.01 (m, 2H, **3**), 3.90 – 3.88 (m, 3H, **4**).

62 1-((1-Methyl-1H-imidazol-5-yl)methyl)-1H-indole-5-carbonitrile

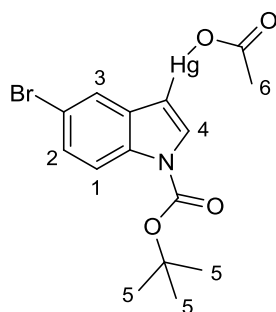
To an ice cold, stirred solution of 1H-indole-5-carbonitrile **56** (195 mg, 1.37 mmol) in 5 mL DMF was added 60% NaH in oil dispersion (131 mg, 3.28 mmol). After hydrogen evolution ceased, **65** (179 mg, 1.07 mmol) was added to stir for 16 hours. After the solvent was removed *in vacuo*, the remainder was partitioned between water and EtOAc (3 x 50 mL). The recombined organic layers were then washed with brine, dried with MgSO_4 and filtered. The product was purified by column chromatography (3:1:96 MeOH:TEA:DCM) and recrystallized from DCM/Hexanes/Toluene to afford pale yellow crystals (110 mg, 34%). $R_f = 0.08$ (1:1:98 MeOH:TEA:DCM). ^1H NMR (300 MHz, CDCl_3) δ 7.98 - 7.97 (m, 1H, ArH), 7.48 - 7.43 (m, 3H, ArH), 7.17 (s, 1H, ArH), 7.10 (d, $J = 3.3$ Hz, 1H, ArH), 5.30 (s, 2H, **6**), 3.34 (s, 3H, **7**).

64 Tert-butyl 5-bromo-1H-indole-1-carboxylate

To a stirring solution of 5-bromo-indole **70** (500 mg, 2.55 mmol) in MeCN (15 mL), was added Boc_2O (612 mg, 2.81 mmol) and a catalytic amount of DMAP. After the reaction was allowed to stir for 16 hours at room temperature, the solvent was removed *in vacuo*. The

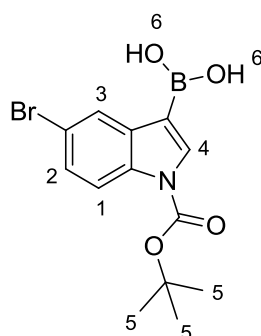
crude was partitioned with water and extracted with EtOAc (3 x 50 mL). The combined organic layers washed with brine, dried with MgSO₄ and filtered. The crude was purified using gravity chromatography (1:9 EtOAc:Hexanes) to yield a clear oil (679 mg, 90%). R_f = 0.30 (1:9 EtOAc:Hexanes). ¹H NMR (400 MHz, CDCl₃) δ 8.02 (d, *J* = 8.9 Hz, 1H, **1**), 7.69 (d, *J* = 2.0 Hz, 1H, **3**), 7.58 (d, *J* = 3.7 Hz, 1H, **4 or 5**), 7.39 (dd, *J* = 8.9, 2.0 Hz, 1H, **2**), 6.50 (d, *J* = 3.7, 1H, **5 or 4**), 1.67 (s, 9H, **6**).

65 Acetoxy(5-bromo-1-(*tert*-butoxycarbonyl)-1*H*-indol-3-yl)mercury



To a stirring solution of **64** (679 mg, 2.29 mmol) in glacial acetic acid (17 mL), was added mercury(II) acetate (729 mg, 2.29 mmol). After the reaction was left to stir overnight, the slurry was filtered to give a solid which was used without further purification (1146 mg, 90%). R_f = 0.50 (20% EtOAc / Hexanes). ¹H NMR (400 MHz, CDCl₃) δ 8.04 (d, *J* = 8.7 Hz, 1H, **1**), 7.69 (d, *J* = 1.9 Hz, 1H, **3**), 7.54 (s, 1H, **4**), 7.41 (dd, *J* = 8.7, 1.9 Hz, 1H, **2**), 2.13 (s, 3H, **6**), 1.67 (s, 9H, **5**).

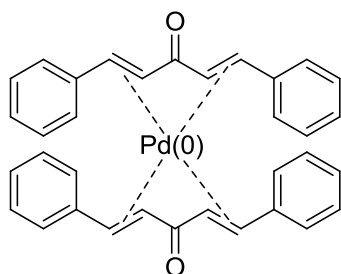
66 (5-Bromo-1-(*tert*-butoxycarbonyl)-1*H*-indol-3-yl)boronic acid



To a solution of **65** (1146 mg, 2.070 mmol) in THF (30 mL) purged with nitrogen was added a solution of borane (155 mg, 20.7 mmol). After the mixture was agitated for 1 hour, water was added (15 mL). The liquid portion was separated from the elemental mercury and concentrated. The residue was dissolved in THF/EtOAc mixture the solids removed by filtration. The filtrate was partitioned between water and brine and the aqueous phases discarded. The organic phase was concentrated *in vacuo* and used without further

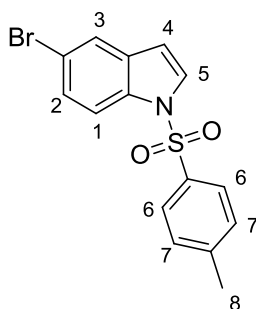
purification (504 mg, 71%). $R_f = 0.16$ (20% EtOAc/ Hexanes). $^1\text{H NMR}$ (300 MHz, DMSO) δ 8.24 – 8.22 (m, 2H, **6**), 8.22 – 8.21 (m, 1H, **Ar**), 8.14 – 8.12 (m, 1H, **Ar**), 8.00 – 7.94 (m, 1H, **Ar**), 7.45 – 7.40 (m, 1H, **Ar**), 1.63 (s, 9H, **5**).

71 Tris(dibenzylideneacetone)dipalladium

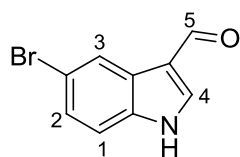


To a solution of hot 50°C MeOH (15 mL) that has been purged with nitrogen, containing dibenzylacetone (460 mg, 1.96 mmol) and sodium acetate (390 mg, 4.75 mmol) was added PdCl_2 (105 mg, 0.592 mmol). The solution was left to stir for 4 hours at 40°C to form a precipitate which was removed by filtration after cooling. After washing with water and acetone, the product was dried *in vacuo* to yield a purple powder (0.365g, 71%).

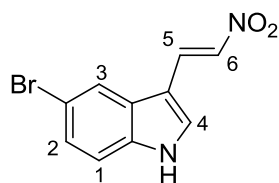
72 5-Bromo-1-tosyl-1H-indole



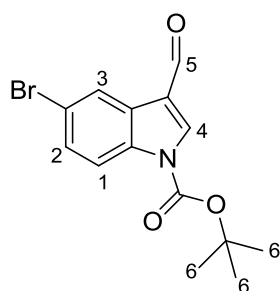
To an ice cold solution of 5-bromoindole **70** (100 mg, 0.511 mmol) in DMF (20 mL) was added 60% NaH in oil dispersion (13.4 mg, 0.561 mmol). After the evolution of hydrogen ceased (15 minutes), tosylchloride (107 mg, 0.561 mmol) was added and the reaction was allowed to proceed for 5 hours. The reaction was quenched by the addition of water (100 mL) and extracted with EtOAc (3 x 50 mL). The recombined organic layers were washed with brine, dried with MgSO_4 , and filtered. Purification with column chromatography (1:9 EtOAc:Hexanes) yielded a white precipitate (95 mg, 53%). $R_f = 0.4$ (1:9 EtOAc:Hexanes). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.88 – 7.84 (m, 1H, **Ar**), 7.75 – 7.74 (m, 1H, **Ar**), 7.74 – 7.71 (m, 1H, **Ar**), 7.67 – 7.64 (m, 1H, **Ar**), 7.57 – 7.55 (m, 1H, **Ar**), 7.42 – 7.37 (m, 1H, **Ar**), 7.25 – 7.23 (m, 1H, **Ar**), 7.23 – 7.20 (m, 1H, **Ar**), 6.60 – 6.57 (m, 1H, **Ar**), 2.35 (s, 3H, **8**).

79 5-Bromo-1H-indole-3-carbaldehyde

To ice cold DMF (9.56 mL) was added POCl₃ (4.66 mL, 51.0 mmol) and allowed to stir for 15 minutes. To this was added 5-bromo-1H-indole **60** (1000 mg, 5.10 mmol) and allowed to stir for 1 hour. The temperature was then increased to 40 °C and allowed to stir for another hour. The reaction was then cooled to RT and then poured onto ice. Aqueous NaOH was then added to pH 6. The reaction was then stirred overnight. The precipitate was then collected by filtration and washed with water. This was then dried under high vacuum to yield a white solid (1010 mg, 88%). ¹H NMR (400 MHz, DMSO) δ 9.92 (s, 1H, **5**), 8.34 (s, 1H, **4**), 8.21 (d, *J* = 1.9 Hz, 1H, **3**), 7.49 (d, *J* = 8.6 Hz, 1H, **1**), 7.40 (dd, *J* = 8.6, 2.0 Hz, 1H, **2**).

78 5-Bromo-3-(2-nitrovinyl)-1H-indole

To a solution of **79** (300 mg, 1.34 mmol) and ammonium acetate (305 mg, 4.02 mmol) was added nitromethane (18 mL) and set to reflux for 1 hour. The solvent was then removed *in vacuo*, washed with ice cold water and extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with brine, dried with MgSO₄ and filtered. The crude was purified by flash chromatography (1:2 EtOAc:Hexanes) to yield an orange powder (201 mg, 56.2%). *R_f* = 0.54 (1:1 EtOAc:Hexanes). ¹H NMR (400 MHz, DMSO) δ 8.44 – 8.33 (m, 4.0 Hz, 1H), 8.29 – 8.24 (m, 1H), 8.23 – 8.18 (m, 1H), 8.13 – 8.02 (m, 1H), 7.55 – 7.43 (m, 1H), 7.42 – 7.33 (m, 1H).

80 Tert-butyl 5-bromo-3-formyl-1*H*-indole-1-carboxylate

To a solution of **79** (200 mg, 0.892 mmol) in THF (40.0 mL) was added Boc_2O (272 mg, 1.25 mmol) and a single crystal of DMAP. After the reaction was allowed to stir for 16 hours at room temperature, the solvent was removed *in vacuo*. The crude was partitioned with water and extracted with EtOAc (3 x 50 mL). The combined organic layers washed with brine, dried with MgSO_4 and filtered. The crude was purified using gravity chromatography (1:4 EtOAc:Hexanes) to yield a clear oil (257 mg, 88.9%). $R_f = 0.80$ (1:1 EtOAc:Hexanes). ^1H NMR (400 MHz, DMSO) δ 10.07 (s, 1H, **5**), 8.72 (s, 1H, **4**), 8.27 (d, $J = 2.0$ Hz, 1H, **3**), 8.06 (d, $J = 8.9$ Hz, 1H, **1**), 7.62 (dd, $J = 8.9, 2.0$ Hz, 1H, **2**), 1.66 (s, 9H, **6**).

Modelling SpdSyn – 2PT9**Protein preparation****Standard dynamics cascade #1**

Minimisation 1		Equilibration	
Algorithm	SD	Steps	10000
Max Steps	5000	Time step	0.001
RMS gradient	0.1	Target temp	300
Energy change	0	Save results frequency	100
Constraints		Adjust velocity frequency	50
Minimisation 2		Production	
Algorithm	CG	Steps	10000
Max Steps	5000	Time step	0.001
RMS gradient	0.0001	Target temp	300
Energy change	0	Temperature coupling decay	5
		Save results frequency	1
Heating		Implicit Solvent model	DDDE
Steps	20000	Electrostatics	Spherical cut-off
Initial temp	0.001		
Target temp	300		
Adjust velocity frequency	50	Fixed atom constrains: all non-hydrogen atoms	
Save results frequency	100		

Standard dynamics cascade #2

Minimisation 1		Equilibration	
Algorithm	SD	Steps	10000
Max Steps	5000	Time step	0.001
RMS gradient	0.1	Target temp	300
Energy change	0	Save results frequency	100
Constraints		Adjust velocity frequency	50
Minimisation 2		Production	
Algorithm	CG	Steps	100
Max Steps	5000	Time step	0.001
RMS gradient	0.0001	Target temp	300
Energy change	0	Temperature coupling decay	5
Constraints		Save results freq	1
		Implicit Solvent model	DDDE
Heating		Electrostatics	Spherical cut-off
Steps	2000		
Initial temp	0.001		
Target temp	300		
Adjust velocity frequency	50	Fixed atom constrains: all non-hydrogen atoms	
Save results frequency	100		

Chapter 8 – Experimental

Solvation

Solvation model	Explicit periodic boundary
Radius of Sphere	20.0
Centre of Mass	
Minimum Distance From Boundary	7.0
Cell Shape	Orthorhombic
Add Counterion	TRUE
Salt Concentration	0.145
Random Seed	314159
Cation Type	Sodium
Anion Type	Chloride

Minimisation

Algorithm	Smart Minimizer
Max Steps	200000
RMS Gradient	0.1
Energy Change	0
Save Results Frequency	0
Implicit Solvent Model	None
Non-bond List Radius	14.0
Non-bond Higher Cut-off Distance	12.0
Non-bond Lower Cut-off Distance	12.0
Electrostatics	Spherical Cut-off
Kappa	0.34
Order	4

Positive control – 4MCHA redock**CDOCKER**

Top Hits	10	Advanced	
Pose Cluster Radius	10	Forcefield	CHARMm
Random conformations	10	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	8
Dynamic target temperature	1000	Ligand Partial Charge Method	Momany-Rone
Include Electrostatic Interactions	True	Random number seed	314159
Orientations to refine		Final Minimisation	Full Potential
Maximum bad orientations	800	Final Minimisation Gradient Tolerance	0
Orientation vdW Energy Threshold	300		
Simulated Annealing	True		
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

Chapter 8 – Experimental

vHTS – Pharmacophore method**Build 3D database**

Conformation method	BEST
Number of conformations	150
Advanced settings	Default

Search 3D database

Input database	Sigma	Max Hits	300
Align ligands	True	Output	All conformations
Search Method	Best	Descriptors	True

CDOCKER

Top Hits	10	Advanced	
Pose Cluster Radius		Forcefield	CHARMm
Random conformations	100	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	
Dynamic target temperature	1000	Ligand Partial Charge	Momany-Rone
Include Electrostatic Interactions	True	Method	
Orientations to refine	20	Random number seed	314159
Maximum bad orientations	800	Final Minimisation	Full Potential
Orientation vdW Energy Threshold	300	Final Minimisation	
Simulated Annealing	True	Gradient Tolerance	0
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

Modelling SpdSyn – 2I7C**Protein preparation****Standard dynamics cascade**

Minimisation 1		Equilibration	
Algorithm	SD	Steps	10000
Max Steps	5000	Time step	0.001
RMS gradient	0.1	Target temp	300
Energy change	0	Adjust velocity frequency	50
Constraints		Save results frequency	100
Minimisation 2		Production	
Algorithm	CG	Steps	1000
Max Steps	5000	Time step	0.001
RMS gradient	0.0001	Target temp	300
Energy change	0	Temperature coupling decay	5
Constraints		Save results frequency	1
Heating		Implicit Solvent model	
Steps	2000		GBSW
Initial temp	0.001	Electrostatics	Spherical cut-off
Target temp	300	Fixed atom constrains: all non-hydrogen atoms	
Adjust velocity frequency	50		
Save results frequency	100		

Solvation

Solvation model	Explicit periodic boundary
Radius of Sphere	20
Centre of Mass	
Minimum Distance From Boundary	7
Cell Shape	Orthorhombic
Add Counterion	TRUE
Salt Concentration	0.145
Random Seed	314159
Cation Type	Sodium
Anion Type	Chloride

Minimisation

Algorithm	Smart Minimizer
Max Steps	40000
RMS Gradient	0
Energy Change	0
Save Results Frequency	0
Implicit Solvent Model	None
Non-bond List Radius	14
Non-bond Higher Cut-off Distance	12
Non-bond Lower Cut-off Distance	12
Electrostatics	Particle Mesh Ewald
Kappa	0.34
Order	4

Chapter 8 – Experimental

CDOCKER

Top Hits	10	Advanced	
Pose Cluster Radius		Forcefield	CHARMm
Random conformations	30	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	8
Dynamic target temperature	1000	Ligand Partial Charge Method	Momany-Rone
Include Electrostatic Interactions	TRUE	Random number seed	314159
Orientations to refine	15	Final Minimisation	Full Potential
Maximum bad orientations	800	Final Minimisation Gradient Tolerance	0
Orientation vdW Energy Threshold	300		
Simulated Annealing	TRUE		
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

Positive control – AdoDATO redock**CDOCKER**

Top Hits	10	Advanced	
Pose Cluster Radius		Forcefield	CHARMm
Random conformations	30	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	8
Dynamic target temperature	1000	Ligand Partial Charge Method	Momany-Rone
Include Electrostatic Interactions	TRUE	Random number seed	314159
Orientations to refine	15	Final Minimisation	Full Potential
Maximum bad orientations	800	Final Minimisation Gradient Tolerance	0
Orientation vdW Energy Threshold	300		
Simulated Annealing	TRUE		
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

Hotspot guided screen - LibDock**LibDock**

Number of hotspots	100	Input	ChemDB
Docking Tolerance	0.25	Conformation method	None
Docking Pre-set	High-Quality	Minimisation algorithm	Do not minimise

CDocking of 1600 LibDock results**CDOCKER**

Top Hits	1	Advanced	
Pose Cluster Radius		Forcefield	CHARMm
Random conformations	45	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	8
Dynamic target temperature	1000	Ligand Partial Charge Method	Momany-Rone
Include Electrostatic Interactions	TRUE	Random number seed	314159
Orientations to refine	15	Final Minimisation	Full Potential
Maximum bad orientations	800	Final Minimisation Gradient Tolerance	0
Orientation vdW Energy Threshold	300		
Simulated Annealing	TRUE		
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

SpdSyn docking parameters for designed compounds**CDOCKER**

Top Hits	10	Advanced	
Pose Cluster Radius		Forcefield	CHARMm
Random conformations	30	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	8
Dynamic target temperature	1000	Ligand Partial Charge Method	Momany-Rone
Include Electrostatic Interactions	TRUE	Random number seed	314159
Orientations to refine	10	Final Minimisation	Full Potential
Maximum bad orientations	800	Final Minimisation Gradient Tolerance	0
Orientation vdW Energy Threshold	300		
Simulated Annealing	TRUE		
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

Modelling PFT – 2R2L**PFT protein preparation****Standard dynamics cascade**

Minimisation 1		Equilibration	
Algorithm	SD	Steps	1000
Max Steps	500	Time step	0.001
RMS gradient	0.1	Target temp	300
Energy change	0	Save results frequency	100
Constraints		Adjust velocity frequency	50
Minimisation 2		Production	
Algorithm	CG	Steps	10000
Max Steps	500	Time step	0.001
RMS gradient	0.0001	Target temp	300
Energy change	0	Temperature coupling decay	5
Constraints		Save results frequency	100
Heating		Implicit Solvent model	
Steps	2000		DDDE
Initial temp	0.001	Electrostatics	Spherical cut-off
Target temp	300	Fixed atom constrains: all non-hydrogen atoms	
Adjust velocity frequency	50		
Save results frequency	100		

Solvation

Solvation model	Explicit periodic boundary
Radius of Sphere	20.0
Centre of Mass	
Minimum Distance From Boundary	7.0
Cell Shape	Orthorhombic
Add Counterion	TRUE
Salt Concentration	0.145
Random Seed	314159
Cation Type	Sodium
Anion Type	Chloride

Minimisation

Algorithm	Smart Minimizer
Max Steps	20000
RMS Gradient	0.1
Energy Change	0
Save Results Frequency	0
Implicit Solvent Model	GB
Nonbonded List Radius	14.0
Nonbonded Higher Cut-off Distance	12.0
Nonbonded Lower Cut-off Distance	12.0
Electrostatics	Spherical Cut-off
Kappa	0.34
Order	4

2R2L docking protocols**CDOCKER**

Top Hits	10	Advanced	
Pose Cluster Radius		Forcefield	CHARMm
Random conformations	16	Use full potential	FALSE
Dynamic Steps	1000	Grid Extension	8
Dynamic target temperature	1000	Ligand Partial Charge Method	Momany- Rone
Include Electrostatic Interactions	TRUE	Random number seed	314159
Orientations to refine	10	Final Minimisation	Full Potential
Maximum bad orientations	800	Final Minimisation Gradient Tolerance	0
Orientation vdW Energy Threshold	300		
Simulated Annealing	TRUE		
Heating Steps	2000		
Heat Target Temperature	700		
Cooling Steps	5000		
Cooling Target Temperature	300		

9. References

1. Bleicher, K.H.; Böhm, H.-.; Müller, K.; Alanine, A.I. Hit and lead generation: Beyond high-throughput screening. *Nature Reviews Drug Discovery* **2003**, *2*, 369-378.
2. Le Marec, C. History of the medicinal opium: From poppy to allialoids of opium. *Douleurs* **2004**, *5*, 83-98.
3. Forbes, T.R. Why is it called 'beautiful lady'? A note on belladonna. *Bulletin of the New York Academy of Medicine: Journal of Urban Health* **1977**, *53*, 403-406.
4. Fisher, H.H. Origin and uses of ipecac. *Economic Botany* **1973**, *27*, 231-234.
5. Tsoucalas, G.; Karamanou, M.; Androutsos, G. Travelling through time with aspirin, a healing companion. *European Journal of Inflammation* **2011**, *9*, 13-16.
6. Fleming, A. Penicillin as a chemotherapeutic agent. *British journal of experimental pathology* **1940**, *10*, 266.
7. Chain, E.; Florey, H.W.; Adelaide, M.B.; Gardner, A.D.; Oxford, D.M.; Heatley, N.G.; Jennings, M.A.; Orr-Ewing, J.; Sanders, A.G. Penicillin as a chemotherapeutic agent. *The Lancet* **1940**, *236*, 226-228.
8. Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to improve RD productivity: The pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery* **2010**, *9*, 203-214.
9. Food and Drug Administration Code of Federal Regulations - § **314.108 New drug product exclusivity. 1999**,
10. Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery* **2004**, *3*, 711-715.
11. Kolchinsky, P. *The entrepreneurs guide to a Biotech startup*. Peter Kolchinsky: 2004;
12. Drews, J. Drug discovery: A historical perspective. *Science* **2000**, *287*, 1960-1964.
13. Maxwell, S.R.J.; Webb, D.J. Receptor functions. *Medicine* **2008**, *36*, 344-349.
14. Soudijn, W.; Van Wijngaarden, I.; Ijzerman, A.P. Allosteric modulation of G-protein coupled receptors: Perspectives and recent developments. *Drug Discovery Today* **2004**, *9*, 752-758.
15. Ren, H.; Yang, B.F.; Rainov, N.G. Receptor tyrosine kinases as therapeutic targets in malignant glioma. *Reviews on Recent Clinical Trials* **2007**, *2*, 87-101.
16. Weir, C.J. Ion channels, receptors, agonists and antagonists. *Anaesthesia and Intensive Care Medicine* **2010**, *11*, 377-383.
17. Gashaw, I.; Ellinghaus, P.; Sommer, A.; Asadullah, K. What makes a good drug target? *Drug Discovery Today* **2012**, *17*, S24-S30.
18. Patrick, G.L. *An introduction to medicinal chemistry*. Oxford University Press: Oxford, 2009; pp. 752.
19. Domagk, G.J. Ein Beitrag zur Chemotherapie der bakteriellen infektionen. *Deutsche medizinische Wochenschrift* **1935**, *61*, 250.
20. Singh, J.; Petter, R.C.; Baillie, T.A.; Whitty, A. The resurgence of covalent drugs. *Nature Reviews Drug Discovery* **2011**, *10*, 307-317.
21. Noe, M.C.; Gilbert, A.M. Targeted Covalent Enzyme Inhibitors. *Annual Reports in Medicinal Chemistry* **2012**, *47*, 413-439.

22. Cramer, R.D. The inevitable QSAR renaissance. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 35-38.
23. Wlodawer, A.; Vondrasek, J. Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 249-284.
24. Lauri, G.; Bartlett, P.A. CAVEAT: A program to facilitate the design of organic molecules. *Journal of Computer-Aided Molecular Design* **1994**, *8*, 51-66.
25. Gillet, V.J.; Newell, W.; Mata, P.; Myatt, G.; Sike, S.; Zsoldos, Z.; Johnson, A.P. SPROUT: Recent developments in the de novo design of molecules. *Journal of Chemical Information and Computer Sciences* **1994**, *34*, 207-217.
26. Böhm, H. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *Journal of Computer-Aided Molecular Design* **1992**, *6*, 61-78.
27. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Research* **2000**, *28*, 235-242.
28. Overington, J.P.; Al-Lazikani, B.; Hopkins, A.L. How many drug targets are there? *Nature Reviews Drug Discovery* **2006**, *5*, 993-996.
29. Muegge, I.; Oloff, S. Advances in virtual screening. *Drug Discovery Today: Technologies* **2006**, *3*, 405-411.
30. Sperandio, O.; Petitjean, M.; Tuffery, P. wwLigCSRre: A 3D ligand-based server for hit identification and optimization. *Nucleic Acids Res.* **2009**, *37*, W504-W509.
31. Kalyanamoorthy, S.; Chen, Y.-P. Structure-based drug design to augment hit discovery. *Drug Discovery Today* **2011**, *16*, 831-839.
32. Badrinarayan, P.; Sastry, G.N. Virtual high throughput screening in new lead identification. *Combinatorial Chemistry and High Throughput Screening* **2011**, *14*, 840-860.
33. Irwin, J.J.; Shoichet, B.K. ZINC - A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling* **2005**, *45*, 177-182.
34. Wermuth, C.G. Selective optimization of side activities: The SOSA approach. *Drug Discovery Today* **2006**, *11*, 160-164.
35. Raju, T.N. The Nobel chronicles. 1988: James Whyte Black, (b 1924), Gertrude Elion (1918-99), and George H Hitchings (1905-98). *Lancet* **2000**, *355*, 1022.
36. Yung-Chi, C.; Prusoff, W.H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I_{50}) of an enzymatic reaction. *Biochemical Pharmacology* **1973**, *22*, 3099-3108.
37. Williams, D.H.; Cox, J.P.L.; Doig, A.J.; Gardner, M.; Gerhard, U.; Kaye, P.T.; Lal, A.R.; Nicholls, I.A.; Salter, C.J.; Mitchell, R.C. Toward the semiquantitative estimation of binding constants. Guides for peptide-peptide binding in aqueous solution. *Journal of the American Chemical Society* **1991**, *113*, 7020-7030.
38. Bissantz, C.; Kuhn, B.; Stahl, M. A medicinal chemist's guide to molecular interactions. *Journal of Medicinal Chemistry* **2010**, *53*, 5061-5084.
39. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **2001**, *46*, 3-26.
40. Veber, D.F.; Johnson, S.R.; Cheng, H.-.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular properties that influence

- the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry* **2002**, *45*, 2615-2623.
41. Shah, P.; Jogani, V.; Bagchi, T.; Misra, A. Role of Caco-2 Cell Monolayers in Prediction of Intestinal Drug Absorption. *Biotechnology Progress* **2006**, *22*, 186-198.
42. Abraham, M.H.; Chadha, H.S.; Mitchell, R.C. Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain. *Journal of Pharmaceutical Sciences* **1994**, *83*, 1257-1268.
43. Egan, W.J.; Lauri, G. Prediction of intestinal permeability. *Advanced Drug Delivery Reviews* **2002**, *54*, 273-289.
44. Giaginis, C.; Theocharis, S.; Tsantili-Kakoulidou, A. Current toxicological aspects on drug and chemical transport and metabolism across the human placental barrier. *Expert Opinion on Drug Metabolism and Toxicology* **2012**, *8*, 1263-1275.
45. Holcberg, G.; Tsadkin-Tamir, M.; Sapir, O.; Huleihel, M.; Mazor, M.; Zvi, Z.B. New aspects in placental drug transfer. *Israel Medical Association Journal* **2003**, *5*, 873-876.
46. Nebert, D.W.; Russell, D.W. Clinical importance of the cytochromes P450. *Lancet* **2002**, *360*, 1155-1162.
47. Anzenbacher, P.; Anzenbacherová, E. Cytochromes P450 and metabolism of xenobiotics. *Cellular and Molecular Life Sciences* **2001**, *58*, 737-747.
48. Sotaniemi, E.A.; Arranto, A.J.; Pelkonen, O.; Pasanen, M. Age and cytochrome P450-linked drug metabolism in humans: An analysis of 226 subjects with equal histopathologic conditions. *Clinical Pharmacology and Therapeutics* **1997**, *61*, 331-339.
49. McGraw, J.; Waller, D. Cytochrome P450 variations in different ethnic populations. *Expert Opinion on Drug Metabolism and Toxicology* **2012**, *8*, 371-382.
50. Lampe, J.W.; King, I.B.; Li, S.; Grate, M.T.; Barale, K.V.; Chen, C.; Feng, Z.; Potter, J.D. Brassica vegetables increase and apiaceous vegetables decrease cytochrome P450 1A2 activity in humans: Changes in caffeine metabolite ratios in response to controlled vegetable diets. *Carcinogenesis* **2000**, *21*, 1157-1162.
51. Morgan, E.T. Regulation of cytochromes P450 during inflammation and infection. *Drug Metabolism Reviews* **1997**, *29*, 1129-1188.
52. Roden, D.M.; George Jr., A.L. The genetic basis of variability in drug responses. *Nature Reviews Drug Discovery* **2002**, *1*, 37-44.
53. Jakoby, W.B.; Ziegler, D.M. The enzymes of detoxication. *Journal of Biological Chemistry* **1990**, *265*, 20715-20718.
54. Vane, J.R. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nature: New biology* **1971**, *231*, 232-235.
55. Nielsen, N.M.; Bundgaard, H. Evaluation of glycolamide esters and various other esters of aspirin as true aspirin prodrugs. *Journal of Medicinal Chemistry* **1989**, *32*, 727-734.
56. Casi, G.; Neri, D. Antibody-drug conjugates: Basic concepts, examples and future perspectives. *Journal of Controlled Release* **2012**, *161*, 422-428.
57. Frazier, J.L.; Han, J.E.; Lim, M.; Olivi, A. Immunotherapy Combined with Chemotherapy in the Treatment of Tumors. *Neurosurgery Clinics of North America* **2010**, *21*, 187-194.
58. Bodor, N.; Buchwald, P. Soft drug design: General principles and recent applications. *Medical Research Reviews* **2000**, *20*, 58-101.

Chapter 9 – References

59. Stańczak, A.; Ferra, A. Prodrugs and soft drugs. *Pharmacological Reports* **2006**, *58*, 599-613.
60. Mortelmans, K.; Zeiger, E. The Ames Salmonella/microsome mutagenicity assay. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* **2000**, *455*, 29-60.
61. Warmke, J.W.; Ganetzky, B. A family of potassium channel genes related to eag in Drosophila and mammals. *Proceedings of the National Academy of Sciences* **1994**, *91*, 3438-3442.
62. Meanwell, N.A. Synopsis of some recent tactical application of bioisosteres in drug design. *Journal of Medicinal Chemistry* **2011**, *54*, 2529-2591.
63. WORLD HEALTH ORGANIZATION WHO Expert Committee on Malaria: Twentieth Report. *Technical Report* **2000**, *892*,
64. Tripathi, R.P.; Mishra, R.C.; Dwivedi, N.; Tewari, N.; Verma, S.S. Current status of malaria control. *Current Medicinal Chemistry* **2005**, *12*, 2643-2659.
65. Teklehaimanot, A.; Mejia, P. Malaria and poverty. *Annals of the New York Academy of Sciences* **2008**, *1136*, 32-37.
66. Worrall, E.; Basu, S.; Hanson, K. Is malaria a disease of poverty? A review of the literature. *Tropical Medicine and International Health* **2005**, *10*, 1047-1059.
67. Wells, T.N.C.; Alonso, P.L.; Gutteridge, W.E. New medicines to improve control and contribute to the eradication of malaria. *Nature Reviews Drug Discovery* **2009**, *8*, 879-891.
68. Egan, T.J. Haemozoin (malaria pigment): A unique crystalline drug target. *Drug Discovery Today: TARGETS* **2003**, *2*, 115-124.
69. O'Neill, P.M.; Ward, S.A.; Berry, N.G.; Jeyadevan, J.P.; Biagini, G.A.; Asadollaly, E.; Park, B.K.; Bray, P.G. A medicinal chemistry perspective on 4-aminoquinoline antimalarial drugs. *Current Topics in Medicinal Chemistry* **2006**, *6*, 479-507.
70. Russel, P.F. *Man's Mastery of Malaria*. Oxford University Press: Oxford, 1955;
71. Ginsburg, H. Should chloroquine be laid to rest? *Acta Trop.* **2005**, *96*, 16-23.
72. Pradines, B.; Mamfoumbi, M.M.; Tall, A.; Sokhna, C.; Koeck, J.-.; Fusai, T.; Mosnier, J.; Czarnecki, E.; Spiegel, A.; Trape, J.-.; Kombila, M.; Rogier, C. In vitro activity of tafenoquine against the asexual blood stages of *Plasmodium falciparum* isolates from Gabon, Senegal, and Djibouti. *Antimicrobial Agents and Chemotherapy* **2006**, *50*, 3225-3226.
73. Fidock, D.A.; Nomura, T.; Talley, A.K.; Cooper, R.A.; Dzekunov, S.M.; Ferdig, M.T.; Ursos, L.M.B.; Bir Singh Sidhu, A.; Naudé, B.; Deitsch, K.W.; Su, X.; Wootton, J.C.; Roepe, P.D.; Wellems, T.E. Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Molecular Cell* **2000**, *6*, 861-871.
74. Bray, P.G.; Martin, R.E.; Tilley, L.; Ward, S.A.; Kirk, K.; Fidock, D.A. Defining the role of PfCRT in Plasmodium falciparum chloroquine resistance. *Molecular Microbiology* **2005**, *56*, 323-333.
75. Lakshmanan, V.; Bray, P.G.; Verdier-Pinard, D.; Johnson, D.J.; Horrocks, P.; Muhle, R.A.; Alakpa, G.E.; Hughes, R.H.; Ward, S.A.; Krogstad, D.J.; Sidhu, A.B.S.; Fidock, D.A. A critical role for PfCRT K76T in *Plasmodium falciparum* verapamil-reversible chloroquine resistance. *EMBO J.* **2005**, *24*, 2294-2305.
76. Fidock, D.A.; Nomura, T.; Cooper, R.A.; Su, X.; Talley, A.K.; Wellems, T.E. Allelic modifications of the cg2 and cg1 genes do not alter the chloroquine response of drug-resistant *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* **2000**, *110*, 1-10.

77. Hiebsch, R.R.; Raub, T.J.; Wattenberg, B.W. Primaquine blocks transport by inhibiting the formation of functional transport vesicles: Studies in a cell-free assay of protein transport through the Golgi apparatus. *Journal of Biological Chemistry* **1991**, *266*, 20323-20328.
78. Famin, O.; Ginsburg, H. Differential effects of 4-aminoquinoline-containing antimalarial drugs on hemoglobin digestion in *Plasmodium falciparum*-infected erythrocytes. *Biochemical Pharmacology* **2002**, *63*, 393-398.
79. Hoppe, H.C.; Van Schalkwyk, D.A.; Wiehart, U.I.M.; Meredith, S.A.; Egan, J.; Weber, B.W. Antimalarial quinolines and artemisinin inhibit endocytosis in *Plasmodium falciparum*. *Antimicrobial Agents and Chemotherapy* **2004**, *48*, 2370-2378.
80. Reed, M.B.; Sallba, K.J.; Caruana, S.R.; Kirk, K.; Cowman, A.F. Pgh1 modulates sensitivity and resistance to multiple antimalarials in *Plasmodium falciparum*. *Nature* **2000**, *403*, 906-909.
81. Johnson, D.J.; Fidock, D.A.; Mungthin, M.; Lakshmanan, V.; Sidhu, A.B.S.; Bray, P.G.; Ward, S.A. Evidence for a central role for PfCRT in conferring *Plasmodium falciparum* resistance to diverse antimalarial agents. *Molecular Cell* **2004**, *15*, 867-877.
82. Duraisingh, M.T.; Cowman, A.F. Contribution of the *pfmdr1* gene to antimalarial drug-resistance. *Acta Tropical* **2005**, *94*, 181-190.
83. Wongsrichanalai, C.; Pickard, A.L.; Wernsdorfer, W.H.; Meshnick, S.R. Epidemiology of drug-resistant malaria. *Lancet Infectious Diseases* **2002**, *2*, 209-218.
84. Meshnick, S.R.; Thomas, A.; Ranz, A.; Xu, C.-.; Pan, H.-. Artemisinin (qinghaosu): The role of intracellular hemin in its mechanism of antimalarial action. *Molecular and Biochemical Parasitology* **1991**, *49*, 181-189.
85. Posner, G.H.; Oh, C.H. A regiospecifically oxygen-18 labeled 1,2,4-trioxane: A simple chemical model system to probe the mechanism(s) for the antimalarial activity of artemisinin (Qinghaosu). *Journal of the American Chemical Society* **1992**, *114*, 8328-8329.
86. Posner, G.H.; O'Neill, P.M. Knowledge of the proposed chemical mechanism of action and cytochrome P450 metabolism of antimalarial trioxanes like artemisinin allows rational design of new antimalarial peroxides. *Accounts of Chemical Research* **2004**, *37*, 397-404.
87. Haynes, R.K.; Cheu, K.-.; Chan, H.-.; Wong, H.-.; Li, K.-.; Tang, M.M.-.; Chen, M.-.; Guo, Z.-.; Guo, Z.-.; Sinniah, K.; Witte, A.B.; Coghi, P.; Monti, D. Interactions between Artemisinins and other Antimalarial Drugs in Relation to the Cofactor Model-A Unifying Proposal for Drug Action. *ChemMedChem* **2012**, *7*, 2204-2226.
88. World Health Organization Guidelines for the treatment of malaria. **2010**,
89. Dondorp, A.M.; Nosten, F.; Yi, P.; Das, D.; Physo, A.P.; Tarning, J.; Lwin, K.M.; Ariey, F.; Hanpithakpong, W.; Lee, S.J.; Ringwald, P.; Silamut, K.; Imwong, M.; Chotivanich, K.; Lim, P.; Herdman, T.; An, S.S.; Yeung, S.; Singhasivanon, P.; Day, N.P.J.; Lindegardh, N.; Socheat, D.; White, N.J. Artemisinin resistance in *Plasmodium falciparum* malaria. *New England Journal of Medicine* **2009**, *361*, 455-467.
90. Krungkrai, J.; Imprasittichai, W.; Otjungreed, S.; Pongsabut, S.; Krungkrai, S.R. Artemisinin resistance or tolerance in human malaria patients. *Asian Pacific Journal of Tropical Medicine* **2010**, *3*, 748-753.
91. Breman, J.G. Resistance to artemisinin-based combination therapy. *The Lancet Infectious Diseases* **2012**, *12*, 820-822.

92. Krishna, S.; Uhlemann, A.-.; Haynes, R.K. Artemisinins: Mechanisms of action and potential for resistance. *Drug Resistance Updates* **2004**, *7*, 233-244.
93. Price, R.N.; Uhlemann, A.-.; Brockman, A.; McGready, R.; Ashley, E.; Phaipun, L.; Patel, R.; Laing, K.; Looareesuwan, S.; White, N.J.; Nosten, F.; Krishna, S. Mefloquine resistance in *Plasmodium falciparum* and increased pfmdr1 gene copy number. *Lancet* **2004**, *364*, 438-447.
94. Falco, E.A.; Goodwin, L.G.; Hitchings, G.H.; Rollo, I.M.; Russell, P.B. 2:4-diaminopyrimidines- a new series of antimalarials. *British journal of pharmacology and chemotherapy* **1951**, *6*, 185-200.
95. Gregson, A.; Plowe, C.V. Mechanisms of resistance of malaria parasites to antifolates. *Pharmacological Reviews* **2005**, *57*, 117-145.
96. Triglia, T.; Cowman, A.F. The mechanism of resistance to sulfa drugs in *Plasmodium falciparum*. *Drug Resistance Updates* **1999**, *2*, 15-19.
97. Fry, M.; Pudney, M. Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4'-chlorophenyl) cyclohexyl]-3-hydroxy-1,4-naphthoquinone (566C80). *Biochemical Pharmacology*. **1992**, *43*, 1545-1553.
98. Srivastava, I.K.; Rottenberg, H.; Vaidya, A.B. Atovaquone, a broad spectrum antiparasitic drug, collapses mitochondrial membrane potential in a malarial parasite. *Journal of Biological Chemistry* **1997**, *272*, 3961-3966.
99. Srivastava, I.K.; Morrley, J.M.; Darrouzet, E.; Daldal, F.; Vaidya, A.B. Resistance mutations reveal the atovaquone-binding domain of cytochrome b in malaria parasites. *Molecular Microbiology* **1999**, *33*, 704-711.
100. Fisher, N.; Majid, R.A.; Antoine, T.; Al-Helal, M.; Warman, A.J.; Johnson, D.J.; Lawrenson, A.S.; Ranson, H.; O'Neill, P.M.; Ward, S.A.; Biagini, G.A. Cytochrome b mutation Y268S conferring atovaquone resistance phenotype in malaria parasite results in reduced parasite bc 1 catalytic turnover and protein expression. *Journal of Biological Chemistry* **2012**, *287*, 9731-9741.
101. Geigenmuller, U.; Nierhaus, K.H. Tetracycline can inhibit tRNA binding to the ribosomal P site as well as to the A site. *European Journal of Biochemistry* **1986**, *161*, 723-726.
102. Alam, A.; Goyal, M.; Iqbal, M.S.; Pal, C.; Dey, S.; Bindu, S.; Maity, P.; Bandyopadhyay, U. Novel antimalarial drug targets: Hope for new antimalarial drugs. *Expert Review of Clinical Pharmacology* **2009**, *2*, 469-489.
103. Maruyoshi, K.; Nonaka, K.; Sagane, T.; Demura, T.; Yamaguchi, T.; Matsumori, N.; Oishi, T.; Murata, M. Conformational change of spermidine upon interaction with adenosine triphosphate in aqueous solution. *Chemistry - A European Journal* **2009**, *15*, 1618-1626.
104. Wright, R.K.; Buehler, B.A.; Schott, S.N.; Rennert, O.M. Spermine and spermidine, modulators of the cell surface enzyme adenylate cyclase. *Pediatric Research* **1978**, *12*, 830-833.
105. Beninati, S.; Gentile, V.; Caraglia, M.; Lentini, A.; Tagliaferri, P.; Abbruzzese, A. Tissue transglutaminase expression affects hypusine metabolism in BALB/c 3T3 cells. *FEBS Letters* **1998**, *437*, 34-38.
106. Williams, K. Interactions of polyamines with ion channels. *Biochemical Journal* **1997**, *325*, 289-297.
107. Nichols, C.G.; Lopatin, A.N. Inward rectifier potassium channels. *Annual Review of Physiology* **1997**, *59*, 171-191.
108. Igarashi, K.; Kashiwagi, K. Polyamines: Mysterious modulators of cellular functions *Biochemical and*

- Biophysical Research Communications* **2000**, 271, 559-564.
109. Kaiser, A.; Gottwald, A.; Maier, W.; Seitz, H.M. Targeting enzymes involved in spermidine metabolism of parasitic protozoa - A possible new strategy for anti-parasitic treatment. *Parasitology Research* **2003**, 91, 508-516.
110. Pegg, A.E. S-Adenosylmethionine decarboxylase. *Essays in Biochemistry* **2009**, 46, 25-45.
111. Pegg, A.E. S-adenosylmethionine decarboxylase: a brief review. *Cell Biochemistry and Function* **1984**, 2, 11-15.
112. Müller, S.; Da'dara, A.; Lüersen, K.; Wrenger, C.; Das Gupta, R.; Madhubala, R.; Walter, R.D. In the human malaria parasite *Plasmodium falciparum*, polyamines are synthesized by a bifunctional ornithine decarboxylase, S-adenosylmethionine decarboxylase. *Journal of Biological Chemistry* **2000**, 275, 8097-8102.
113. Pegg, A.E. Recent advances in the biochemistry of polyamines in eukaryocytes. *Biochemical Journal* **1986**, 234, 249-262.
114. Haider, N.; Eschbach, M.-.; De Souza Dias, S.; Gilberger, T.-.; Walter, R.D.; Lüersen, K. The spermidine synthase of the malaria parasite *Plasmodium falciparum*: Molecular and biochemical characterisation of the polyamine synthesis enzyme. *Molecular Biochemical Parasitology* **2005**, 142, 224-236.
115. Pegg, A.E.; Coward, J.K. Growth of mammalian cells in the absence of the accumulation of spermine. *Biochemical and Biophysical Research Communications* **1985**, 133, 82-89.
116. Nishimura, K.; Nakatsu, F.; Kashiwagi, K.; Ohno, H.; Saito, T.; Igarashi, K. Essential role of S-adenosylmethionine decarboxylase in mouse embryonic development. *Genes to Cells* **2002**, 7, 41-47.
117. Pendeville, H.; Carpino, N.; Marine, J.; Takahashi, Y.; Muller, M.; Martial, J.A.; Cleveland, J.L. The ornithine decarboxylase gene is essential for cell survival during early murine development. *Molecular and Cellular Biology* **2001**, 21, 6549-6558.
118. Guo, K.; Chang, W.; Newell, P.C. Isolation of spermidine synthase gene (spsA) of *Dictyostelium discoideum*. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1999**, 1449, 211-216.
119. Gilroy, C.; Olenyik, T.; Roberts, S.C.; Ullman, B. Spermidine synthase is required for virulence of *Leishmania donovani*. *Infectious Immunology* **2011**, 79, 2764-2769.
120. Roberts, S.C.; Jiang, Y.; Jardim, A.; Carter, N.S.; Heby, O.; Ullman, B. Genetic analysis of spermidine synthase from *Leishmania donovani*. *Molecular and Biochemical Parasitology* **2001**, 115, 217-226.
121. Jin, Y.; Bok, J.W.; Guzman-De-Peña, D.; Keller, N.P. Requirement of spermidine for developmental transitions in *Aspergillus nidulans*. *Molecular Microbiology* **2002**, 46, 801-812.
122. Dufe, V.T.; Qiu, W.; Müller, I.B.; Hui, R.; Walter, R.D.; Al-Karadaghi, S. Crystal Structure of *Plasmodium falciparum* Spermidine Synthase in Complex with the Substrate Decarboxylated S-adenosylmethionine and the Potent Inhibitors 4MCHA and AdoDATO. *Journal of Molecular Biology* **2007**, 373, 167-177.
123. Vedadi, M.; Lew, J.; Artz, J.; Amani, M.; Zhao, Y.; Dong, A.; Wasney, G.A.; Gao, M.; Hills, T.; Brokx, S.; Qiu, W.; Sharma, S.; Diassiti, A.; Alam, Z.; Melone, M.; Mulichak, A.; Wernimont, A.; Bray, J.; Loppnau, P.; Plotnikova, O.; Newberry, K.; Sundararajan, E.; Houston, S.; Walker, J.; Tempel, W.; Bochkarev, A.; Koziaradzki, I.; Edwards, A.; Arrowsmith, C.; Roos, D.; Kain, K.; Hui, R. Genome-scale protein expression and structural biology of

- geranylgeranyltransferase. Subunit composition and metal requirements. *Journal of Biological Chemistry* **1992**, *267*, 17438-17443.
140. Yokoyama, K.; Gelb, M.H. Purification of a mammalian protein geranylgeranyltransferase. Formation and catalytic properties of an enzyme-geranylgeranyl pyrophosphate complex. *Journal of Biological Chemistry* **1993**, *268*, 4055-4060.
141. Seabra, M.C.; Brown, M.S.; Slaughter, C.A.; Sudhof, T.C.; Goldstein, J.L. Purification of component A of rab geranylgeranyl transferase: Possible identity with the choroideremia gene product. *Cell* **1992**, *70*, 1049-1057.
142. Seabra, M.C.; Goldstein, J.L.; Sudhof, T.C.; Brown, M.S. Rab geranylgeranyl transferase. A multisubunit enzyme that prenylates GTP-binding proteins terminating in Cys-X-Cys or Cys-Cys. *Journal of Biological Chemistry* **1992**, *267*, 14497-14503.
143. Hast, M.A.; Fletcher, S.; Cummings, C.G.; Pusateri, E.E.; Blaskovich, M.A.; Rivas, K.; Gelb, M.H.; Van Voorhis, W.C.; Sebt, S.M.; Hamilton, A.D.; Beese, L.S. Structural Basis for Binding and Selectivity of Antimalarial and Anticancer Ethylenediamine Inhibitors to Protein Farnesyltransferase. *Chemistry and Biology* **2009**, *16*, 181-192.
144. Park, H.-.; Boduluri, S.R.; Moomaw, J.F.; Casey, P.J.; Beese, L.S. Crystal structure of protein farnesyltransferase at 2.25 angstrom resolution. *Science* **1997**, *275*, 1800-1804.
145. Van Voorhis, W.C.; Rivas, K.L.; Bendale, P.; Nallan, L.; Hornéy, C.; Barrett, L.K.; Bauer, K.D.; Smart, B.P.; Ankala, S.; Hucce, O.; Verlinde, C.L.M.J.; Chakrabarti, D.; Strickland, C.; Yokoyama, K.; Buckner, F.S.; Hamilton, A.D.; Williams, D.K.; Lombardo, L.J.; Floyd, D.; Gelb, M.H. Efficacy, pharmacokinetics, and metabolism of tetrahydroquinoline inhibitors of Plasmodium falciparum protein farnesyltransferase. *Antimicrobial Agents and Chemotherapy* **2007**, *51*, 3659-3671.
146. Rappé, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard III, W.A.; Skiff, W.M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society* **1992**, *114*, 10024-10035.
147. Halgren, T.A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* **1996**, *17*, 490-519.
148. Halgren, T.A. MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. *Journal of Computational Chemistry* **1999**, *20*, 730-748.
149. Maple, J.R.; Hwang, M.-.; Jalkanen, K.J.; Stockfisch, T.P.; Hagler, A.T. Derivation of class II force fields: V. Quantum force field for amides, peptides, and related compounds. *Journal of Computational Chemistry* **1998**, *19*, 430-458.
150. Maple, J.R.; Hwang, M.-.; Stockfisch, T.P.; Dinur, U.; Waldman, M.; Ewig, C.S.; Hagler, A.T. Derivation of class II force fields. I. Methodology and quantum force field for the alkyl functional group and alkane molecules. *Journal of Computational Chemistry* **1994**, *15*, 162-182.
151. Brooks, B.R.; Brooks III, C.L.; Mackerell Jr., A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.W.; Post, C.B.; Pu, J.Z.; Schaefer, M.; Tidor, B.; Venable, R.M.; Woodcock, H.L.; Wu, X.; Yang, W.; York, D.M.; Karplus, M.

- CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry* **2009**, *30*, 1545-1614.
152. Brooks, B.R.; Brucoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **1983**, *4*, 187-217.
153. Jones, J.E. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924**, *106*, 463-477.
154. Mazur, J.; Jernigan, R.L. Distance-dependent dielectric constants and their application to double-helical DNA. *Biopolymers* **1991**, *31*, 1615-1629.
155. Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *Journal of the American Chemical Society* **1990**, *112*, 6127-6129.
156. Nicholls, A.; Honig, B. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *Journal of Computational Chemistry* **1991**, *12*, 435-445.
157. Chu, J.; Trout, B.L.; Brooks, B.R. A super-linear minimization scheme for the nudged elastic band method. *Journal of Chemical Physics* **2003**, *119*, 12708-12717.
158. Powell, M.J.D. Some convergence properties of the conjugate gradient method. *Mathematical Programming* **1976**, *11*, 42-49.
159. Petrova, S.S.; Solov'Ev, A.D. The Origin of the Method of Steepest Descent. *Historia Mathematica* **1997**, *24*, 361-375.
160. Ginsburg, T. The conjugate gradient method. *Numerische Mathematik* **1963**, *5*, 191-200.
161. Cvijovic, D.; Klinowski, J. Taboo search: An approach to the multiple minima problem. *Science* **1995**, *267*, 664-666.
162. Sakae, Y.; Hiroyasu, T.; Miki, M.; Okamoto, Y. New conformational search method using genetic algorithm and knot theory for proteins. *Pacific Symposium on Biocomputing 2011, PSB 2011* **2011**, 217-228.
163. Allen, M.P.; Tildesley, D.J. Periodic Boundary Conditions and Potential Truncation, In *Computer Simulation of Liquids*, Anonymous ; Oxford University Press: Oxford, 1987;
164. Haile, J.M. *Molecular Dynamic Simulations: Elementary Methods*. John Wiley & Sons: New York, 1992;
165. Kirkpatrick, S.; Gelatt Jr., C.D.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671-680.
166. Wu, G.; Robertson, D.H.; Brooks, C.L.; Vieth, M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER: A CHARMM-based MD docking algorithm. *Journal of Computational Chemistry* **2003**, *24*, 1549-1562.
167. Vieth, M.; Hirst, J.D.; Kolinski, A.; Brooks, C.L. Assessing energy functions for flexible docking. *Journal of Computational Chemistry* **1998**, *19*, 1612-1622.
168. Bolton, E.E.; Wang, Y.; Thiessen, P.A.; Bryant, S.H. Chapter 12 PubChem: Integrated Platform of Small Molecules and Biological Activities. *Annual Reports in Computational Chemistry* **2008**, *4*, 217-241.
169. Chen, J.H.; Linstead, E.; Swamidass, S.J.; Wang, D.; Baldi, P. ChemDB update - Full-text search and virtual chemical

- space. *Bioinformatics* **2007**, *23*, 2348-2351.
170. Sigma-Aldrich Product database. 2013,
171. Rao, S.N.; Head, M.S.; Kulkarni, A.; LaLonde, J.M. Validation studies of the site-directed docking program LibDock. *Journal of Chemical Information and Modeling* **2007**, *47*, 2159-2171.
172. Ghoshal, N.; Achari, B.; Ghoshal, T.K. Computer aids in drug design - Highlights. *Polish journal of pharmacology and pharmacy* **1996**, *48*, 359-377.
173. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe Jr., E.W. Computational methods in drug discovery. *Pharmacology Reviews* **2014**, *66*, 334-395.
174. Jain, A.N. Scoring noncovalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities. *Journal of Computer-Aided Molecular Design* **1996**, *10*, 427-440.
175. Krammer, A.; Kirchhoff, P.D.; Jiang, X.; Venkatachalam, C.M.; Waldman, M. LigScore: A novel scoring function for predicting binding affinities. *Journal of Molecular Graphics and Modelling* **2005**, *23*, 395-407.
176. Böhm, H.-. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of Computer-Aided Molecular Design* **1994**, *8*, 243-256.
177. Böhm, H.-. Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *Journal of Computer-Aided Molecular Design* **1998**, *12*, 309-323.
178. Gehlhaar, D.K.; Verkhivker, G.M.; Rejto, P.A.; Sherman, C.J.; Fogel, D.B.; Fogel, L.J.; Freer, S.T. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: Conformationally flexible docking by evolutionary programming. *Chemistry and Biology* **1995**, *2*, 317-324.
179. Workman, P. The opportunities and challenges of personalized genome-based molecular therapies for cancer: Targets, technologies, and molecular chaperones. *Cancer Chemotherapy and Pharmacology, Supplement* **2003**, *52*, S45-S56.
180. Muegge, I.; Martin, Y.C. A general and fast scoring function for protein-ligand interactions: A simplified potential approach. *Journal of Medicinal Chemistry* **1999**, *42*, 791-804.
181. Muegge, I. PMF scoring revisited. *Journal of Medicinal Chemistry* **2006**, *49*, 5895-5902.
182. Spassov, V.Z.; Flook, P.K.; Yan, L. LOOPER: A molecular mechanics-based algorithm for protein loop prediction. *Protein Engineering, Design and Selection* **2008**, *21*, 91-100.
183. Laurie, A.T.R.; Jackson, R.M. Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics* **2005**, *21*, 1908-1916.
184. Greaser, M. High throughput docking for library design and library prioritization. *Proteins: Structure, Function and Genetics* **2001**, *43*, 113-124.
185. Jacobsson, M.; Gäredal, M.; Schultz, J.; Karlén, A. Identification of Plasmodium falciparum spermidine synthase active site binders through structure-based virtual screening. *Journal of Medicinal Chemistry* **2008**, *51*, 2777-2786.
186. Cheng, A.; Merz Jr., K.M. Prediction of aqueous solubility of a diverse set of compounds using quantitative structure-property relationships. *Journal of Medicinal Chemistry* **2003**, *46*, 3572-3580.
187. Sufrin, J.R.; Spiess, A.J.; Kramer, D.L.; Libby, P.R.; Miller, J.T.; Bernacki,

- R.J.; Lee, Y.; Borchardt, R.T.; Porter, C.W. Targeting 5'-deoxy-5'-(methylthio)adenosine phosphorylase by 5'-haloalkyl analogues of 5'-deoxy-5'-(methylthio)adenosine. *Journal of Medicinal Chemistry* **1991**, *34*, 2600-2606.
188. Li, J.; Wei, H.; Zhou, M.-. Structure-guided design of a methyl donor cofactor that controls a viral histone H3 lysine 27 methyltransferase activity. *Journal of Medicinal Chemistry* **2011**, *54*, 7734-7738.
189. Robins, M.J.; Hansske, F.; Wnuk, S.F.; Kanai, T. Nucleic acid related compounds. 66. Improved synthesis of 5'-chloro-5'-deoxy- and 5'-S-aryl(or alkyl)-5'-thionucleosides. *Canadian Journal of Chemistry* **1991**, *69*, 1468-1474.
190. Townsend, A.P.; Roth, S.; Williams, H.E.L.; Stylianou, E.; Thomas, N.R. New S-adenosyl-L-methionine analogues: Synthesis and reactivity studies. *Organic Letters* **2009**, *11*, 2976-2979.
191. Marasco Jr., C.J.; Kramer, D.L.; Miller, J.; Porter, C.W.; Bacchi, C.J.; Rattendi, D.; Kucera, L.; Iyer, N.; Bernacki, R.; Pera, P.; Sufrin, J.R. Synthesis and evaluation of analogues of 5'-[(Z)-4-amino-2-butenyl]methylamino)-5'-deoxyadenosine as inhibitors of tumor cell growth, trypanosomal growth, and HIV-1 infectivity. *Journal of Medicinal Chemistry* **2002**, *45*, 5112-5122.
192. Murai, N.; Yonaga, M.; Tanaka, K. Palladium-catalyzed direct hydroxymethylation of aryl halides and triflates with potassium acetoxymethyltrifluoroborate. *Organic Letters* **2012**, *14*, 1278-1281.
193. Mayor, M.; Scheffold, R.; Walder, L. Synthesis of vitamin B12 derivatives with a peripheral metal binding site. *Helvetica Chimica Acta* **1997**, *80*, 1183-1189.
194. Inman, M.; Moody, C.J. Synthesis of indolequinones from bromoquinones and enamines mediated by Cu(OAc)₂·2H₂O. *Journal of Organic Chemistry* **2010**, *75*, 6023-6026.
195. Nuñez, S.A.; Yeung, K.; Fox, N.S.; Phillips, S.T. A structurally simple self-immolative reagent that provides three distinct, simultaneous responses per detection event. *The Journal of Organic Chemistry* **2011**, *76*, 10099-10113.
196. Quiroz-Florentino, H.; Hernández-Benitez, R.I.; Aviña, J.A.; Burgueño-Tapia, E.; Tamariz, J. Total synthesis of naturally occurring furan compounds 5-[(4-hydroxybenzyl)oxy]methyl}-2-furaldehyde and pichiafuran C. *Synthesis* **2011**, 1106-1112.
197. Tremblay, M.R.; Luu-The, V.; Leblanc, G.; Noël, P.; Breton, E.; Labrie, F.; Poirier, D. Spironolactone-related inhibitors of type II 17 β -hydroxysteroid dehydrogenase: Chemical synthesis, receptor binding affinities, and proliferative/antiproliferative activities. *Bioorganic and Medicinal Chemistry* **1999**, *7*, 1013-1023.
198. Swamy, K.C.K.; Kumar, N.N.B.; Balaraman, E.; Kumar, K.V.P.P. Mitsunobu and Related Reactions: Advances and Applications. *Chemical Reviews* **2009**, *109*, 2551-2651.
199. Robins, M.J.; Peng, Y.; Damaraju, V.L.; Mowles, D.; Barron, G.; Tackaberry, T.; Young, J.D.; Cass, C.E. Improved syntheses of 5'-S-(2-Aminoethyl)-6-N-(4-nitrobenzyl)-5'-thioadenosine (SAENTA), analogues, and fluorescent probe conjugates: Analysis of cell-surface human equilibrative nucleoside transporter 1 (hENT1) levels for prediction of the antitumor efficacy of gemcitabine. *Journal of Medicinal Chemistry* **2010**, *53*, 6040-6053.
200. Kung, P.-.; Zehnder, L.R.; Meng, J.J.; Kupchinsky, S.W.; Skalitzy, D.J.; Johnson, M.C.; Maegley, K.A.; Ekker, A.; Kuhn, L.A.; Rose, P.W.; Bloom, L.A. Design, synthesis, and biological evaluation of novel human 5'-deoxy-5'-methylthioadenosine phosphorylase (MTAP) substrates. *Bioorganic and Medicinal Chemistry Letters* **2005**, *15*, 2829-2833.

201. Pignot, M.; Pljevaljic, G.; Weinhold, E. Efficient synthesis of s-adenosyl-L-homocysteine natural product analogues and their use to elucidate the structural determinant for cofactor binding of the DNA methyltransferase M-HhaI. *European Journal of Organic Chemistry* **2000**, 549-555.
202. Williams, D.B.G.; Lawton, M. Drying of organic solvents: Quantitative evaluation of the efficiency of several desiccants. *The Journal of Organic Chemistry* **2010**, *75*, 8351-8354.
203. von Deyn, W.; York, W.S.; Albersheim, P.; Darvill, A.G. 1-Alkoxyamino-1-deoxy alditols, useful u.v.-absorbing derivatives of neutral and acidic oligosaccharides. *Carbohydrate Research* **1990**, *201*, 135-144.
204. Zhou, X.; Chen, W.; Xu, C.; Fan, S.; Xie, Y.; Zhong, W.; Wang, L.; Li, S. (S)-3-(4-(2-(5-Methyl-2-phenyloxazol-4-yl)ethoxy)phenyl)-2-(piperazin-1-yl)propanoic acid compounds: Synthesis and biological evaluation of dual PPAR α / γ agonists. *Bioorganic and Medicinal Chemistry Letters* **2010**, *20*, 2605-2608.
205. Thayumanavan, S. Invertible amphiphilic polymers. **2007**.
206. Rehman, A.; Soni, A.; Naik, K.; Nair, S.; Palle, V.P.; Dastidar, S.; Ray, A.; Alam, M.S.; Salman, M.; Cliffe, I.A.; Sattigeri, V. Synthesis and biological activity of N-substituted aminocarbonyl-1,3-dioxolanes as VLA-4 antagonists. *Bioorganic and Medicinal Chemistry Letters* **2010**, *20*, 5514-5520.
207. Richter, J.M.; Whitefield, B.W.; Maimone, T.J.; Lin, D.W.; Castroviejo, M.P.; Baran, P.S. Scope and mechanism of direct indole and pyrrole couplings adjacent to carbonyl compounds: Total synthesis of acremoxin A and oxazinin 3. *Journal of the American Chemical Society* **2007**, *129*, 12857-12869.
208. Raffier, L.; Piva, O. Application of the diastereoselective photodeconjugation of α,β -unsaturated esters to the synthesis of gymnastatin H. *Beilstein Journal of Organic Chemistry* **2011**, *7*, 151-155.
209. Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V.V.; Noodleman, L.; Sharpless, K.B.; Fokin, V.V. Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. *Journal of the American Chemical Society* **2005**, *127*, 210-216.
210. Berg, R.; Straub, B.F. Advancements in the mechanistic understanding of the copper-catalyzed azide-alkyne cycloaddition. *Beilstein Journal of Organic Chemistry* **2013**, *9*, 2715-2750.
211. Tahtaoui, C.; Parrot, I.; Klotz, P.; Guillier, F.; Galzi, J.-.; Hibert, M.; Ilien, B. Fluorescent pirenzepine derivatives as potential bitopic ligands of the human M1 muscarinic receptor. *Journal of Medicinal Chemistry* **2004**, *47*, 4300-4315.
212. Schröder, T.; Gartner, M.; Grab, T.; Bräse, S. A new azide staining reagent based on "click chemistry". *Organic and Biomolecular Chemistry* **2007**, *5*, 2767-2769.
213. Cegielska, B.; Kacprzak, K.M. Simple and convenient protocol for staining of organic azides on TLC plates by ninhydrin. A new application of an old reagent. *Chemia Analityczna* **2009**, *54*, 807-812.
214. Kotsuki, H.; Sugino, A.; Sakai, H.; Yasuoka, H. A novel synthesis of chiral DBU/DBN-related molecules for use in asymmetric base catalysis. *Heterocycles* **2000**, *53*, 2561-2567.
215. Patil, V.; Guerrant, W.; Chen, P.C.; Gryder, B.; Benicewicz, D.B.; Khan, S.I.; Tekwani, B.L.; Oyelere, A.K. Antimalarial and antileishmanial activities of histone deacetylase inhibitors with triazole-linked cap group. *Bioorganic and Medicinal Chemistry* **2010**, *18*, 415-425.
216. Smilkstein, M.; Sriwilajaroen, N.; Kelly, J.X.; Wilairat, P.; Riscoe, M. Simple and Inexpensive Fluorescence-Based

- Technique for High-Throughput Antimalarial Drug Screening. *Antimicrobial Agents and Chemotherapy* **2004**, *48*, 1803-1806.
217. Peng, H.; Carrico, D.; Thai, V.; Blaskovich, M.; Bucher, C.; Pusateri, E.E.; Sebti, S.M.; Hamilton, A.D. Synthesis and evaluation of potent, highly-selective, 3-aryl-piperazinone inhibitors of protein geranylgeranyltransferase-I. *Organic and Biomolecular Chemistry* **2006**, *4*, 1768-1784.
218. Glenn, M.P.; Chang, S.-.; Hornéy, C.; Rivas, K.; Yokoyama, K.; Pusateri, E.E.; Fletcher, S.; Cummings, C.G.; Buckner, F.S.; Pendyala, P.R.; Chakrabarti, D.; Sebti, S.M.; Gelb, M.; Van Voorhis, W.C.; Hamilton, A.D. Structurally simple, potent, Plasmodium selective farnesyltransferase inhibitors that arrest the growth of malaria parasites. *Journal of Medicinal Chemistry* **2006**, *49*, 5710-5727.
219. Makler, M.T.; Ries, J.M.; Williams, J.A.; Bancroft, J.E.; Piper, R.C.; Gibbins, B.L.; Hinrichs, D.J. Parasite lactate dehydrogenase as an assay for Plasmodium falciparum drug sensitivity. *The American Journal of Tropical Medicine and Hygiene* **1993**, *48*, 739-741.
220. Asoh, K.; Kohchi, M.; Hyoudoh, I.; Ohtsuka, T.; Masubuchi, M.; Kawasaki, K.; Ebiike, H.; Shiratori, Y.; Fukami, T.A.; Kondoh, O.; Tsukaguchi, T.; Ishii, N.; Aoki, Y.; Shimma, N.; Sakaitani, M. Synthesis and structure-activity relationships of novel benzofuran farnesyltransferase inhibitors. *Bioorganic and Medicinal Chemistry Letters* **2009**, *19*, 1753-1757.
221. DeGraw, A.J.; Hast, M.A.; Xu, J.; Mullen, D.; Beese, L.S.; Barany, G.; Distefano, M.D. Caged protein prenyltransferase substrates: Tools for understanding protein prenylation. *Chemical Biology and Drug Design* **2008**, *72*, 171-181.
222. Njoroge, F.G.; Vibulbhan, B.; Pinto, P.; Strickland, C.; Bishop, W.R.; Nomeir, A.; Girijavallabhan, V. Enhanced FTase activity achieved via piperazine interaction with catalytic zinc. *Bioorganic and Medicinal Chemistry Letters* **2006**, *16*, 984-988.
223. Scott Reid, T.; Terry, K.L.; Casey, P.J.; Beese, L.S. Crystallographic analysis of CaaX prenyltransferases complexed with substrates defines rules of protein substrate selectivity. *Journal of Molecular Biology* **2004**, *343*, 417-433.
224. Thoma, J.A.; Koshland Jr., D.E. Competitive inhibition by substrate during enzyme action. Evidence for the induced-fit theory. *Journal of the American Chemical Society* **1960**, *82*, 3329-3333.
225. Miyaura, N.; Yamada, K.; Suzuki, A. A new stereospecific cross-coupling by the palladium-catalyzed reaction of 1-alkenylboranes with 1-alkenyl or 1-alkynyl halides. *Tetrahedron Letters* **1979**, *20*, 3437-3440.
226. Xiong, W.; Yang, C.; Jiang, B. Synthesis of novel analogues of marine indole alkaloids: Mono(indolyl)-4-trifluoromethylpyridines and bis(indolyl)-4-trifluoromethylpyridines as potential anticancer agents. *Bioorganic and Medicinal Chemistry* **2001**, *9*, 1773-1780.
227. Dimroth, O. Directe Einführung von Quecksilber in aromatische Verbindungen. *Berichte der deutschen chemischen Gesellschaft* **1898**, *31*, 2154-2156.
228. Furuya, T.; Strom, A.E.; Ritter, T. Silver-mediated fluorination of functionalized aryl stannaries. *Journal of the American Chemical Society* **2009**, *131*, 1662-1663.
229. Ukai, T.; Kawazura, H.; Ishii, Y.; Bonnet, J.J.; Ibers, J.A. Chemistry of dibenzylideneacetone-palladium(0) complexes. I. Novel tris(dibenzylideneacetone)dipalladium(solvent) complexes and their reactions with quinones. *Journal of Organometallic Chemistry* **1974**, *65*, 253-266.

230. Fresneda, P.M.; Molina, P.; Delgado, S.; Bleda, J.A. Synthetic studies towards the 2-aminopyrimidine alkaloids variolins and meridianins from marine origin. *Tetrahedron Letters* **2000**, *41*, 4777-4780.

231. Vilsmeier, A.; Haack, A. Über die Einwirkung von Halogenphosphor auf Alkyl-formanilide. Eine neue Methode zur Darstellung sekundärer und tertiärer p-Alkylamino-benzaldehyde. *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **1927**, *60*, 119-122.

232. Schumacher, R.W.; Davidson, B.S. Synthesis of didemnolines A-D, N9-substituted β -carboline alkaloids from the marine ascidian *Didemnum* sp. *Tetrahedron* **1999**, *55*, 935-942.

233. Muratore, M.E.; Holloway, C.A.; Pilling, A.W.; Storer, R.I.; Trevitt, G.; Dixon, D.J. Enantioselective Brønsted acid-catalyzed *N*-acyliminium cyclization cascades. *Journal of the American Chemical Society* **2009**, *131*, 10796-10797.

234. Still, W.C.; Kahn, M.; Mitra, A. Rapid chromatographic technique for preparative separations with moderate resolution. *Journal of organic chemistry* **1978**, *43*, 2923-2925.

