Biocatalytic preparation and characterization of alternative substrate of MshB, a mycothiol pathway enzyme

by
Ndivhuwo Olga Muneri

Thesis presented in partial fulfilment of the requirements for the degree Master of Science in Biochemistry at the University of Stellenbosch

Supervisor: Prof Erick Strauss
Faculty of Science

December 2012

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.

Date: December 2012

Research outputs

Article published:

Lamprecht DA, Muneri NO, Eastwood H, Naidoo KJ, Strauss E, Jardine A. An enzyme-initiated Smiles rearrangement enables the development of an assay of MshB, the GlcNAc-Ins deacetylase of mycothiol biosynthesis.
Org. Biomol. Chem. 2012, **10**(27):5278-5288.
Cited: 1

Manuscript in preparation:

Muneri NO, Lamprecht DA, Moracci M and Strauss E. Engineering and characterization of an α-*N*-acetylglucosaminidase for biocatalytic preparation of MshB substrates.

Conference output (poster):

Muneri NO, Lamprecht DA, Eastwood H, Naidoo KJ, Strauss E, Jardine A. Assay development studies of MshB, the GlcNAc-Ins deacetylase of mycothiol biosynthesis Presented at *Trends in Enzymology 2012* (TinE2012) conference, June 3-6, 2012. Georg-August University, Göttingen, Germany

Acknowledgements

- I want to thank my supervisor, Prof. Erick Strauss for giving me the chance to further my studies with his research group after a very challenging undergraduate study. Thank you for your assistance and guidance over the past two years.
- Thanks to all the Strauss Lab members for making the lab a friendly place to work in.
- A special thanks to Dirk, Lizbé, Albert and Bertus for assisting me with my lab work and for the encouraging words.
- Thanks to Dr. Marco Moracci and his lab members, especially Andrea for making me feel at home in the lab, when I was a thousand miles away from my country. I really learned a lot and enjoyed myself during the time I was there.
- A very special thanks to my family and friends for their endless support through it all.
- I also want to thank Suzanne and Matteo for allowing the Prof. to work on my thesis during the most painful and exciting moment in your life. You made my thesis to be very special, you were the red cherry on top of a carrot cake making it to be perfect.
- Finaly, I want to give all my thanks and glory to my Heavenly Father. None of this would have been possible without Him.

Additional Acknowledgements

- The University of Stellenbosch and Prof. E. Strauss for financial support and the opportunity to study at this institution
- The National Research Foundation (NRF) and Institute of Protein Biochemistry, Naples of the National Research Council (NRC) of Italy for financial support
- Mrs. Elsa Malherbe and Dr. Marietjie Stander from the Central Analytical Facility of Stellenbosch University for NMR and LC-MS analyses

To the unknown driver who hit me and drove away during my last year of my Masters degree, I am alive.

Abstract

Mycobacterium tuberculosis (M. tuberculosis), the causative agent of tuberculosis, utilizes mycothiol (MSH) as the major low molecular weight thiol to protect itself against oxidative stress and thereby to ensure its growth and survival. MSH is a pseudo-disaccharide molecule that contains an α(1→1) glycosidic bond, and is biosynthesised in five enzymatic steps involving the enzymes MshA, MshA2, MshB, MshC and MshD. Owing to the essentiality of MSH to *M. tuberculosis*, various studies have focused on the MSH biosynthetic and other MSH-dependent enzymes viewed as potential drug targets for the development of antituberculosis agents. In the course of this study two practical challenges affecting the development of inhibitors of one the MSH biosynthesis pathway enzyme, MshB, were addressed. These challenges entail the lack of a high-throughput continuous assay to determine MshB activity, and the poor availability of the natural and alternative MshB substrates. In this study an alternate MshB substrate was characterized and shown to undergo a rearrangement reaction upon deactylation, which allowed the development of a new continuous assay for MshB activity that uses DNTB (Ellman's reagent). In addition, three new α-thioglycoligases were created from the α-*N*-acetylglucosaminidase of *Clostridium perfringens*. These enzymes showed potential as biocatalysts that can be used for the enzymatic synthesis of thioglycoside-based alternative substrates of MshB.

Opsomming

Mycobacterium tuberculosis, die organisme wat tuberkulose veroorsaak, maak gebruik van mikotiol (MSH) om homself te beskerm teen oksidatiewe stres en sodoende sy groei en oorlewing te verseker. MSH is ’n pseudo-disakkaried molekule met ’n α(1→1) glikosidiese binding, en word in vyf ensimatiese stappe gebiosintetiseer deur die ensieme MshA, MshA2, MshB, MshC en MshD. Weens die noodsaaklikheid van MSH vir *M. tuberculosis* het verskeie vorige studies gefokus op die MSH biosintetiese en ander MSH-afhanklike ensieme as potensiële teikens vir die ontwikkeling van antituberkulose middels. In die loop van hierdie studie is twee praktiese uitdagings wat die ontwikkeling van inhibitors van een van die MSH biosintetiese ensieme, MshB, bemoeilik, aangespreek. Hierdie uitdagings behels die gebrek aan 'n geskikte hoë-deurvloei kontinue essaï vir MshB aktiwiteit, en die lae beskikbaarheid van die natuurlike en alternatiewe MshB substrate. In hierdie studie is alternatiewe MshB substrate gekarakteriseer en is daar gewys dat dit ’n herrangskikkingsreaksie ondergaan na deasetilering, wat die ontwikkeling van ’n nuwe kontinue essaï vir MshB aktiwiteit wat gebruik maak van DTNB (Ellman se reagens), moontlik gemaak het. Verder is drie nuwe α-tioglikoligases ontwikkel van die ensiem α-*N*-asetielglukosaminidase van *Clostridium perfringens*. Hierdie ensieme toon potensiaal as biokataliste wat gebruik kan word in die ensiematiese sintese van tioglikosied-gebaseerde alternatiewe substrate van MshB.

Table of Contents

Chapter 1

Chapter 2

Chapter 4

List of Abbreviations:

λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
^{1}H NMR	proton nuclear magnetic resonance spectroscopy
1-L-Ins-1-P	1-L-myo-inositol-1-phosphate
6xHis-tag	$(His)_6$ peptide tag
AcCoA	acetyl-coenzyme A
AcCySmB	*N*-acetylcysteine-monobromobimane conjugate
AccQ-Fluor™	6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate
ADP	adenosine 5'-diphosphate
AMP	adenine 5'-monophosphate
ATP	adenine 5'-triphosphate
Asp	aspartic acid
BSA	bovine serum albumin
$BF_3.Et_2O$	boron trifluoride diethyl etherate
CD	circular dichrosim
CH_2Cl_2	dichloromethane
CoA	coenzyme A
*Cp*GH89	α-*N*-acetylglucosaminidase of *Clostridium perfringens*
CSA	5'-*O*-[*N*-L-cysteinyl)sulfamonyl]adenosine
CT	C-terminal
C. perfringens	*Clostridium perfringens*
Cys	cysteine

Cys-GlcN-Ins	1-D-*myo*-inositol-2-(L-cysteinyl)-amido-2-deoxy-α-D-glucopyranoside
DNA	deoxyribonucleic acid
DMSO	dimethylsulfoxide
DMPU	1,3-dimethyl-3,4,5,6-tetra-hydro-2(1H)-pyrimidinone
DNTB	5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent)
*d*NP-α-GlcNAc	2,4-dinitrophenyl α-*N*-acetyl-D-glucosaminide
DS-TB	drug-susceptible *M. tuberculosis*
DTT	dithiothreitol
E. coli	*Escherichia coli*
Et_3N	triethylamine
FSA	fluorescamine
GlcNAc	*N*-acetyl-glucosamine
GlcNAc-SBn	benzyl-2-acetamido-2-deoxy-1-thio-3,4,6-tri-*O*-acetyl-α-D-glucopyranose
GlcNAc-SDNP	2,4-dinitrophenyl 2-acetamido-2-deoxy-1-thio-α-D glucopyranose
GlcNAc-SPNP	4-nitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose
GlcNAc-Ins	1-D-myo-inosityl-2-acetamido-2-deoxy-α-D-glycopyranoside
GlcNAc-Ins-3-P	1-D-myo-inosityl-2-acetamido-2-deoxy-α-D-glycopyranoside-3-phosphate
GlcN	glucosamine
GlcN-Ins	1-D-myo-inosityl-2-deoxy-α-D-glycopyranoside
Glu	glutamic acid
Gly	glycine
GNAT	GCN5-related *N*-acetlytransferase
GSH	glutathione

h	hour
H_2O	water
HCl	hydrochloric acid
HEPES	*N*-2-hydroxyethylpiperazine-*N*-2-ethanesulfonic acid
His	histidine
HIV	human immunodeficiency virus
hrs.	hours
HPLC	high pressure liquid chromatography
H_2SO_4	sulfuric acid
Ile	isoleucine
IPTG	isopropyl-β-D-thiogalactoside
IMAC	immobilized metal affinity chromatography
Ins	*myo*-inositol
KCl	potassium chloride
kDa	kilodalton
KOH	potassium hydroxide
KSAc	potassium thioacetate
LB	Luria Bertani
LCMS	liquid chromatography mass spectrometry
Leu	leucine
L-Ins-1P(I1P)	L-*myo*-inositol-1-phosphate
LMW	low molecular weight
Lys (L)	lysine
mBBr	monobromobimane
Mca	mycothiol-S-conjugate amidase
MDR-TB	multidrug-resistant *M. tuberculosis*

$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulfate
min	minute
M. tuberculosis	*Mycobacterium tuberculosis*
Mtr	mycothiol disulfide reductase
MSH	mycothiol
MsmB	bimane derivative
M. smegmatis	*Mycobacterium smegmatis*
MS	mass spectrometry
MscR	mycothiol-*S*-nitrosoreductase/-formaldehyde dehydrogenase
MSSM	mycothiol disulfide
M_r	relative molecular mass
NaCl	sodium chloride / salt / brine
N-AcCys	*N*-acetyl-cysteine
NADH	reduced nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NaOH	sodium hydroxide
NaN_3	sodium azide
NMR	nuclear magnetic resonance spectroscopy
NT	N-terminal
NTB	2-nitro-5-thiobenzoate
OD_{600}	optical density at 600nm
*o*NP-α-GlcNAc	*o*-nitrophenyl α-*N*-acetyl- D -glucosaminide
PCR	polymerase chain reaction
*p*NP-α-GlcNAc	*p*-nitrophenyl α-*N*-acetyl-D-glucosaminide
PPi	pyrophosphate

RE	restriction enzyme
R_f	retention factor
RNA	ribonucleic acid
RNS	reactive nitrogen species
ROS	reactive oxygen species
RS	reactive species
RT	room temperature
s	seconds
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SH	thiol group
SOE PCR	splicing by overlap extension polymerase chain reaction
TAE	triethanolamine
TB	pulmonary tuberculosis
TCEP	*tris*(2-carboxyethyl)phosphine
TLC	thin layer chromatography
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
Trp	tryptophan
$T[SH]_2$	trypanothione
Tyr	tyrosine
U	units (enzyme concentration)
UDP	uridine-diphosphate
UDP-GlcNAc	uridine-diphosphate-*N*-acetyl-glucosamine
UV	ultraviolet
$^w/_w$	mass fraction – weight/weight
$^w/_v$	mass fraction – weight/volume
Val	valine

$^{v}/_{v}$	mass fraction – volume/volume
XDR-TB	extensively drug-resistant *M. tuberculosis*

Chapter 1: Introduction

1.1 Introduction

Tuberculosis (TB) is a lethal infectious disease that claims over a million lives per year worldwide [1-3]. The causative agent of TB is *Mycobacterium tuberculosis*, a pathogenic actinomycete that was discovered by Robert Koch in 1882 [2]. Although infection of a host system by *M. tuberculosis* can result in an active infection, the pathogen is found in a latent form (inactive and non-infectious) in one third of the world's population. This latent infection can develop into active disease as a result of several factors such as unhygienic living conditions, diabetes, anti-tumour necrosis factor therapy or co-infection with the human immunodeficiency virus (HIV). The majority of TB cases are reported in developing countries, owing to poor health care facilities and a high number of HIV infections [1].

The first anti-TB drug, streptomycin, was discovered in the 1940s by Schatz and Waksman [4]. This was followed by the development of other anti-TB drugs including isoniazid, rifamycin, ethambutol and pyrazinamide, among others [2, 4]. Isoniazid and rifamycin are currently the most effective first-line drugs used to treat drug-susceptible TB (DS-TB). This treatment involves a four-drug regimen; whereby isoniazid, rifamycin, ethambutol and pyrazinamide are used for the first 2-3 months, followed by the co-administration of isoniazid and rifamycin for the next 4-6 months [2, 3]. Approximately 95% of DS-TB infections are curable, provided that the treatment is completed and that the right dosage is taken [2].

Incomplete treatment and misuse of drugs (especially pertaining to isoniazid and rifamycin) lead to the development of multi-drug-resistant TB (MDR-TB) [2, 5]. To treat people infected with MDR-TB second-line drugs are required, which includes capreomycin, amikacin and kanamycin. [2, 5]. MDR-TB strain requires a more protracted treatment than DS-TB strain, which involves eight to ten-drug regimens. Improper use of the second-line drugs or interaction of the anti-TB drugs with other drugs, such as antiviral and diabetic drugs, lead to the development of another strain of *M. tuberculosis* known as extensively-drug-resistant TB (XDR-TB). This strain is non-susceptible to both the first- and second-line anti-TB drugs [2]. In addition, it has the highest mortality rate of the different strains.

The rise in the occurrence of the drug-resistant strains and the failure of the current drug regimens to control this rise, highlights the urgent need for the identification of new *M. tuberculosis* drug targets for the development of new treatments against TB. In the next section one of these possible drug targets, namely the disruption of redox balance, is highlighted.

1.2 Redox balance as a drug target

The word “redox” describes two fundamental processes that occur within all living organisms, namely reduction and oxidation reactions [6, 7]. Reduction is defined as a gain of electrons by the oxidising agent (oxidant), while oxidation is the loss of electrons by the reducing agent (reductant) [8, 9]. Maintenance of cellular redox balance is crucial to all living organisms [6, 7]. The redox balance is preserved by the generation and elimination of reactive species (RS) that includes reactive oxygen species (ROS) such as superoxide, hydroxyl, hydrogen peroxide, and reactive nitrogen species (RNS) such as nitric oxide and peroxynitrite.

Upon infection of the human system by *M. tuberculosis* the pathogen is introduced to the innate or adaptive immune mechanism of the host, wherein macrophage activating factors, such as gamma-interferon and tumor necrosis factor-alpha, are released to activate the macrophage. This in turn releases the RS [1, 10]. The RS disrupt the redox balance system of the mycobacterium which results in a condition known as oxidative stress–a state in which more reactive species are generated than can be eliminated from the cell system [7, 9, 11]. The high concentration of RS leads to DNA damage, protein degradation and lipid peroxidation which results in growth inhibition or death of the mycobacterium [9, 11, 12]. Therefore, disturbance of the redox balance system in this pathogenic microorganism can serve as a potential target for the development of new antimicrobials.

Most organisms use low molecular weight thiol-containing compounds to cope with endogenous and exogenous stresses and also to maintain the cellular redox balance [13-15]. These compounds are intracellular redox-regulating antioxidants that act both as buffer and reducing factors to maintain the intracellular reducing environment. A variety of thiol molecules have been identified in various organisms, including the cysteinyl *pseudo*-disaccharide mycothiol (MSH), trypanothione ($T[SH]_2$), coenzyme A (CoASH) and the tripeptide glutathione (GSH) (Figure 1.1). Of these four molecules GSH, which is the major thiol-containing molecule found in

eukaryotes (i.e. humans) and Gram-negative bacteria, is the best studied [14]. In contrast, selected Gram-positive bacteria, trypanosomatids and actinomycetes exploit CoASH, ($T[SH]_2$) and MSH respectively as their low-molecular weight thiols. MSH was first isolated by Newton and co-workers in 1993 from a *Streptomycetes* cell extract [16] and has since been established as the major and most utilized low-molecular weight thiol in the actinomycetes, which includes *M. tuberculosis* and *Mycobacterium smegmatis* [17]. Due to its central role in maintaining the redox balance in *M. tuberculosis*, MSH has been identified as a possible drug target for antituberculosis drug development.

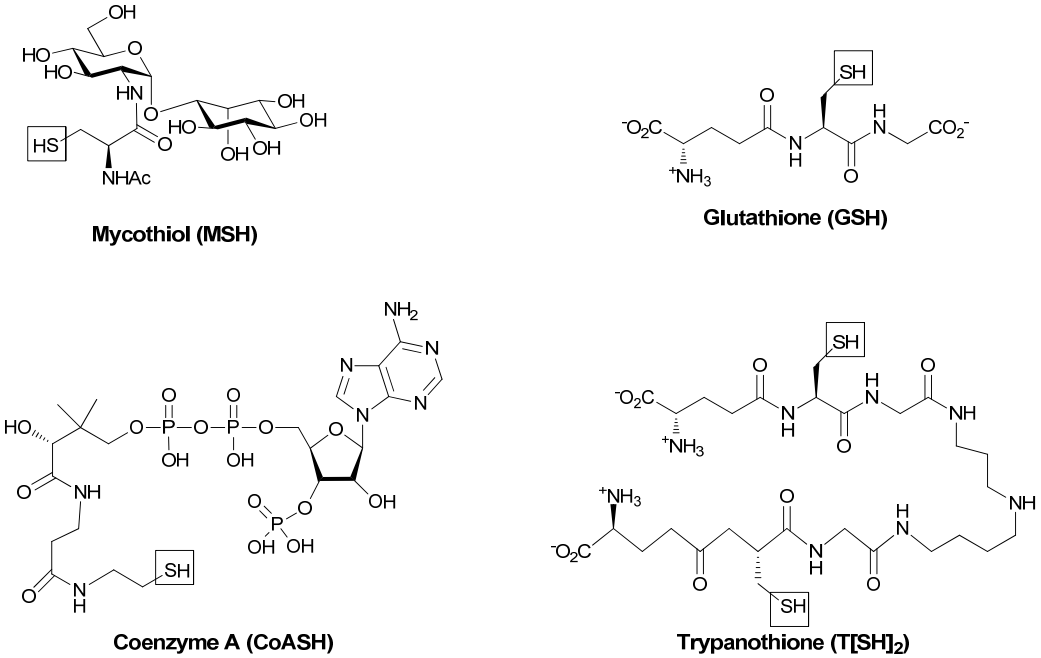


Fig. 1.1: Examples of different low molecular weight thiols used by various organisms to maintain their intracellular redox balance. The thiol group (Shown rectangle) facilitates the maintenance of redox balance within the cell [14].

1.3 Mycothiol as the low molecular weight thiol of *Mycobacterium tuberculosis*

MSH consists of two sugar moieties, *N*-glucosamine and inositol, connected to each other by an α(1→1) glycosidic bond (Figure 1.1) [5, 14, 18]. Additionally, the amine *N*-glucosamine is linked to a cysteine (Cys) via an amide bond, while the amine group of the Cys is acetylated [5, 15].

M. tuberculosis produces MSH in millimolar quantities, which can be explained by the fact that the most important function of MSH in mycobacteria is to maintain the

intracellular redox environment [14, 18-20]. This process is homologous to the role of GSH in both eukaryotes and Gram-negative bacteria. Furthermore mycobacteria utilize MSH to detoxify xenobiotics (alkylating agents and formaldehyde) and MSH can act as a reservoir of cysteine and glucosamine inositol (GlcN-Ins), both of which serves as biosynthetic intermediates and as energy sources during stress.

1.3.1 The mycothiol biosynthesis pathway

MSH is produced in a biosynthetic pathway that involves five enzyme catalysed steps [14, 18, 20]. The first step in the MSH biosynthetic pathway involves the transfer of *N*-acetylglucosamine from UDP-*N*-acetylglucosamine (UDP-GlcNAc) to 1-L-myo-inositol-1-phosphate (1-L-Ins-1-P) producing 3-phospho-1-D-myo-inosityl-2-acetamido-2-deoxy-α-D-glucopyranoside (GlcNAc-Ins-3-P) (Figure 1.2). This step is catalyzed by glycosyltransferase (MshA). GlcNAc-Ins-3-P is dephosphorylated by the phosphatase (MshA2) to yield GlcNAc-Ins. This is followed by deacetylation by GlcNAc-Ins N-deacetylase (MshB) to yield the free amino-sugar, glucosamine-inositol (GlcN-Ins). Subsequently GlcN-Ins is condensed with Cys, at the cost of one ATP molecule, by MSH ligase (MshC), yielding cysteine- glucosamine-inositol (Cys-GlcN-Ins). The final step in the pathway is catalyzed by MSH synthase (MshD), an acetylase that transfers an acetyl group from acetyl-CoA to the free amine group of Cys-GlcN-Ins to produce MSH.

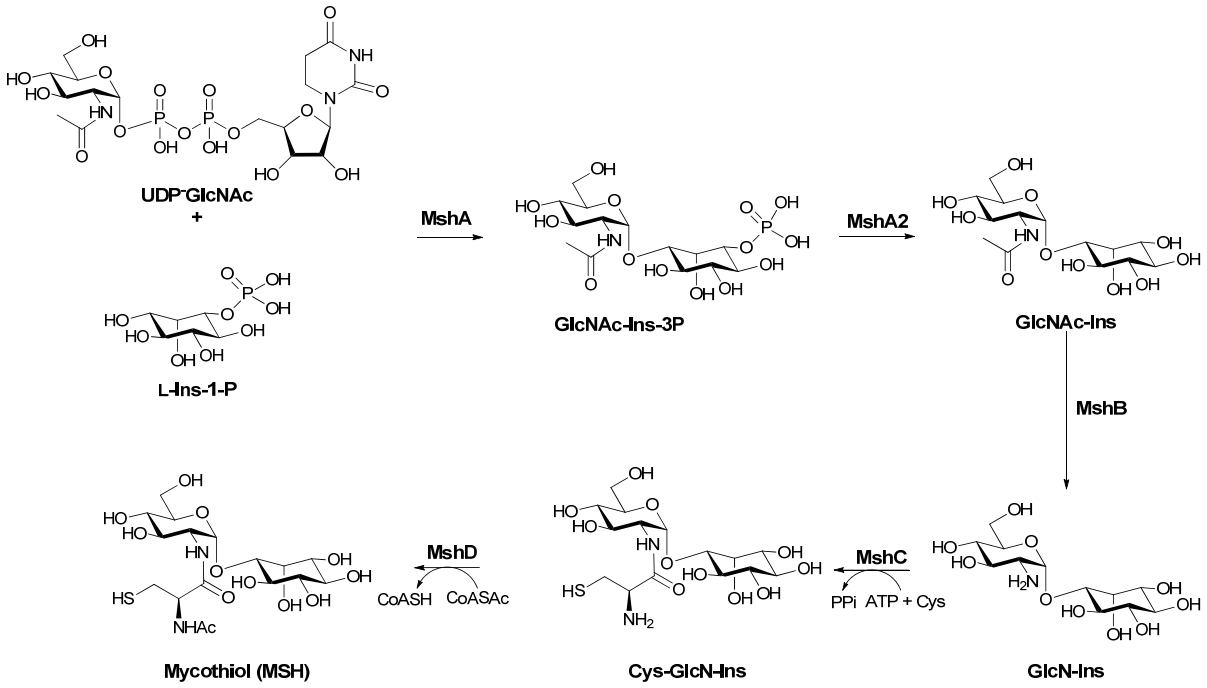
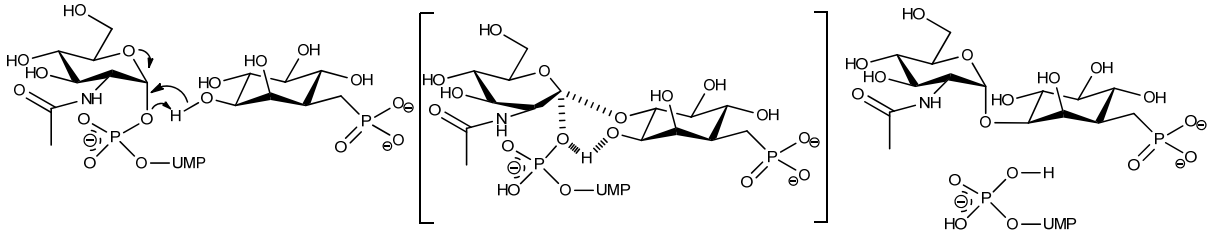


Fig. 1.2: The MSH biosynthesis pathway [14, 18, 20].

1.3.1.1 MshA and MshA2

MshA, a retaining glycotransferase, catalyses the transfer of 1-L-Ins-1-P to α-UDP-GlcNAc to form GlcNAc-(α1,1)-1-D-Ins-3P which has the same anomeric configuration as UDP-GlcNAc (Figure 1.2) [14, 18]. The biochemistry of MshA was first determined by Newton *et al.* (2006) [21] using free cell-extract of *M. smegmatis*, wherein they discovered α-UDP-GlcNAc as the substrate donor and 1-L-Ins-1-P as the substrate acceptor. However, the first crystal structure of MshA was determined by Vetting *et al*. in 2008 [22] of the *Corynebacterium glutamicum* homologue, wherein they utilised two type of crystal forms: $P3_1$ form (APO-enzyme form), binary complex of UDP, and I422 form (ternary complex), binary complex of UDP and 1-L-Ins-1-P. MshA was found to consist of two domains: N-terminal domain and C-terminal domain which display the β/α/β Rossmann-fold type of GT-B fold superfamily with the active site flanked by the domains. Furthermore, MshA was proposed to use an S_Ni substrate-assisted mechanism due to the lack of a nucleophile at the active site, wherein the binding of UDP-GlcNAc leads to the binding of 1-L-Ins-1-P, upon which the hydroxyl group of 1-L-Ins-1-P and β-phosphate of UDP-GlcNAc both act as nucleophiles in the reaction to produce GlcNAc-Ins-3-P (Scheme 1.1) [18, 22]. Subsequently, MshA2 dephosphorylates GlcNAc-Ins-3-P to give GlcNAc-Ins[14, 21] . The identity of MshA2 still remains to be discovered.



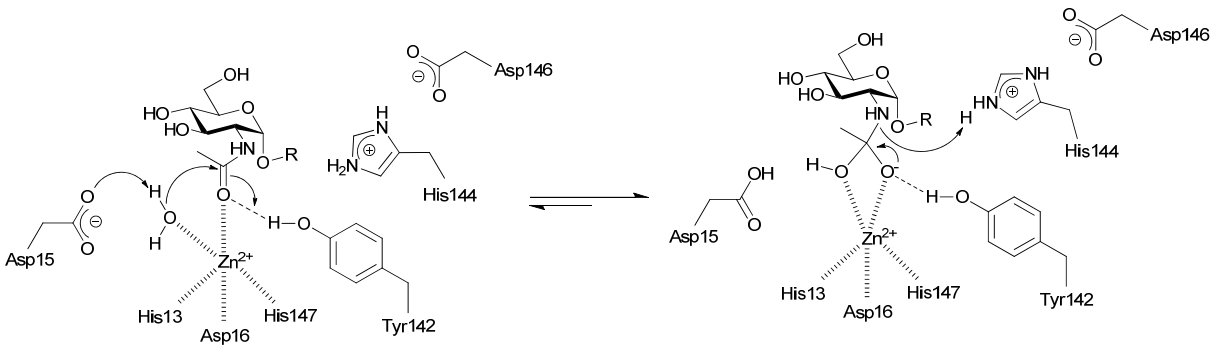
Scheme 1.1: The proposed S_Ni substrate-assisted mechanism of MshA. Scheme adapted from Vetting *et al.* 2008 [22].

1.3.1.2 MshB

MshB is a metalloprotein that catalyses the third reaction of the MSH biosynthetic pathway [14, 23, 24]. It was identified as the key enzyme in the pathway by Newton et al. in 2000 [25] through the discovery of a homologous metalloprotein known as mycothiol *S*-conjugate amidase (Mca). MshB has a high specificity for its natural substrate GlcNAc-Ins, a key intermediate within the MSH biosynthetic pathway. Inactivation studies of the *mshB* gene done by Rawat *et al.* [26] confirmed MshB as

the key enzyme in the MSH biosynthetic pathway. In addition, MshB was demonstrated to be non-essential for the growth or synthesis of MSH in *M. tuberculosis*, since the mutant strain still produce some MSH. This was attributed to its homologous metalloprotein, Mca (Mca will be discussed in more detail in later section).

In the catalytically active site of MshB a Zn^{2+} cation is coordinated to His^{13}, His^{147}, Aps^{16} and two water molecules. The presence of the metal ions was confirmed by two independent crystallization studies [27, 28]. The first study by Maynes *et al.* in 2003 proposed that the catalytic mechanism of MshB is an acid-base mechanism. They proposed that His^{144} acts as the acid and Asp^{15} as the base [27]. This mechanism was confirmed by a study conducted in 2012 by Huang and Herick, in which they discovered an additional amino acid, Tyr^{142}, which facilitate the stability of the oxanion tetrahedral intermediate (Scheme 1.2) [24]. The binding of MshB to its substrate permits the carbonyl oxygen of the acetyl group to replace the water-Zn^{2+} adduct to make way for another water molecule to enter and attack the carbonyl carbon of the acetyl group. The attack on the carbonyl carbon by the water molecule is facilitated by the carboxylate base of Asp^{15}, while His^{144} assists with the dissociation of the intermediate to give GlcN-Ins.

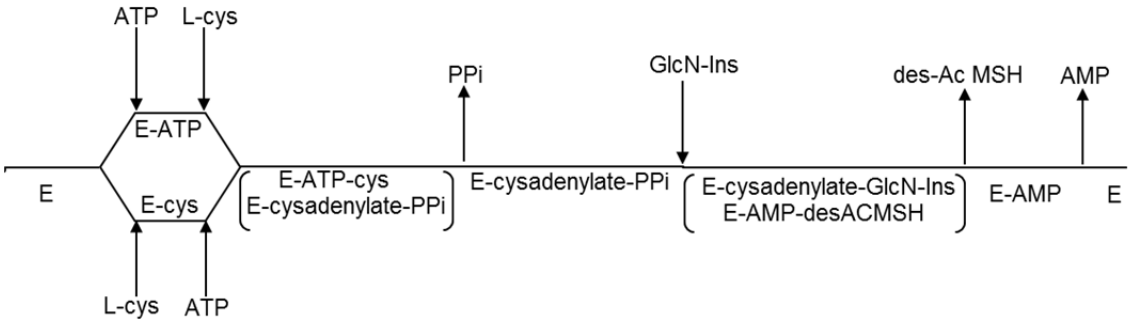


Scheme 1.2: The catalytic mechanism of MshB as proposed by Huang and Herick [24].

1.3.1.3 MshC

MshC is a ligase enzyme that ligates cysteine and GlcN-Ins, the MshB reaction product, at expense of an ATP molecule [18]. This reaction was discovered in 2002 by Sareen *et al.* [29] in a crude extract of *M. smegmatis*. To date only the MshC of *M.*

smegmatis has been expressed heterologously from *E. coli* and purified to homogeneity. Inhibition studies done by Fan *et al*. [30] using 5'-*O*-[*N*-L-cysteinyl)sulfamonyl]adenosine (CSA) as the inhibiting compound, showed that the ligation reaction of Cys to GlcN-Ins occurs in two steps: first the adenylation of Cys followed by the condensation of the Cys to GlcN-Ins. They determined the catalytic mechanism of MshC as a Bi Uni Uni Bi Ping Pong mechanism through steady-state kinetics and position isotope exchange studies (Scheme 1.3). The binding of ATP and Cys to MshC causes the liberation of pyrophosphate (PPi) and the formation of the adenylated-Cys intermediate. After binding of the GlcN-Ins to the enzyme, the Cys is transferred from the enzyme-AMP complex to the GlcN-Ins and Cys-GlcN-Ins and AMP are released from the enzyme.



Scheme 1.3: The catalytic mechanism of MshC. Scheme adapted from Fan *et al.*, 2009 [18].

In an unrelated study, MshC was crystallised bound to CSA revealing a cluster of electron density residues, a flexible loop structure and a zinc ion bound to the thiolate of the CSA in the active site [31]. The fact that the zinc ion was bound to the thiolate of the CSA was taken as evidence that this interaction serves as the basis for the amino acid discriminating ability of the protein.

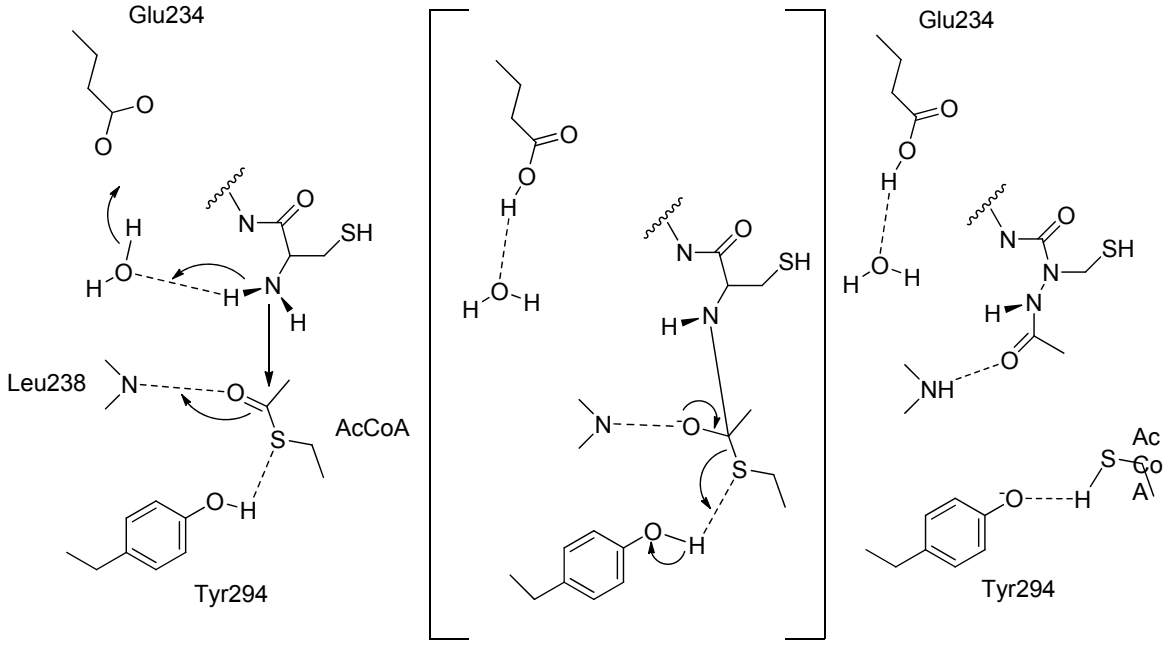
1.3.1.4 MshD

MshD is an acetyltransferase protein which belongs to the GCN5-related *N*-acetlytransferase (GNATs) family [23, 32]. This enzyme catalyse the final reaction in the MSH biosynthetic pathway; transferring an acetyl group to Cys-GlcN-Ins to give MSH (Figure 1.2).

The crystal structure of *M. tuberculosis* MshD has been solved by Vetting *et al.* [33]. The enzyme has been crystallised in a binary complex system with acetyl-CoA

(AcCoA) and as a ternary complex with CoA and desacetylmycothiol (DAM) [18, 33, 34]. MshD contains two GNAT domains, one on the *N*-terminal side and one on the *C*-terminal side. These domains are linked by a random coil. The domains utilise two different to bind to the AcCoA. The C-terminal domain possesses the catalytic activity while the *N*-terminal domain stabilizes the structure of the protein.

The catalytic mechanism of *M. tuberculosis* MshD was proposed by Vetting *et a*l. in 2006 [34]. There are three amino acids important for activity: the glutamic acid at position 324 (Glu^{324}), the tyrosine at position 294 (Tyr^{294}) and the leucine at position 238 (Leu^{238}) (Scheme 1.4). Glu^{324} and Tyr^{294} acts as a base and acid respectively, while Leu238 stabilises the tetrahedral intermediate [18, 33, 34]. Glu^{324} coordinates a water molecule and abstracts a proton from the free amine group of Cys-GlcN-Ins. Tyr^{294} protonate the sulfhydryl-group of CoA after the acetyl is transferred to form the final product, MSH.



Scheme 1.4: The catalytic mechanism of MshD. Scheme adapted from Vetting *et al*., 2006 [34].

1.3.2 Function of mycothiol in *M. tuberculosis*

Actinomycetes including *M. tuberculosis* employ MSH in three different metabolic processes: defensive reactions against ROS, RNS and xenobiotic agents; degradation reactions during nutrient starvation; and metabolic reactions wherein

MSH act as a cofactor for the growth of mycobacteria [15]. All of these reactions are catalysed by various MSH-dependent enzymes (Figure 1.3) [14, 15, 35]. The defensive reactions are catalysed by mycothiol disulfide reductase (Mtr), mycothiol-*S*-nitrosoreductase/-formaldehyde dehydrogenase (MscR), thiol peroxidases and Mca. Mca also catalyse the degradation reactions that provide nutrients for the mycobacteria. It is the best studied enzyme of these pathways, as it has been identified and assayed first [14, 23]. The metabolic reactions are catalysed by maleylpyruvate isomerase.

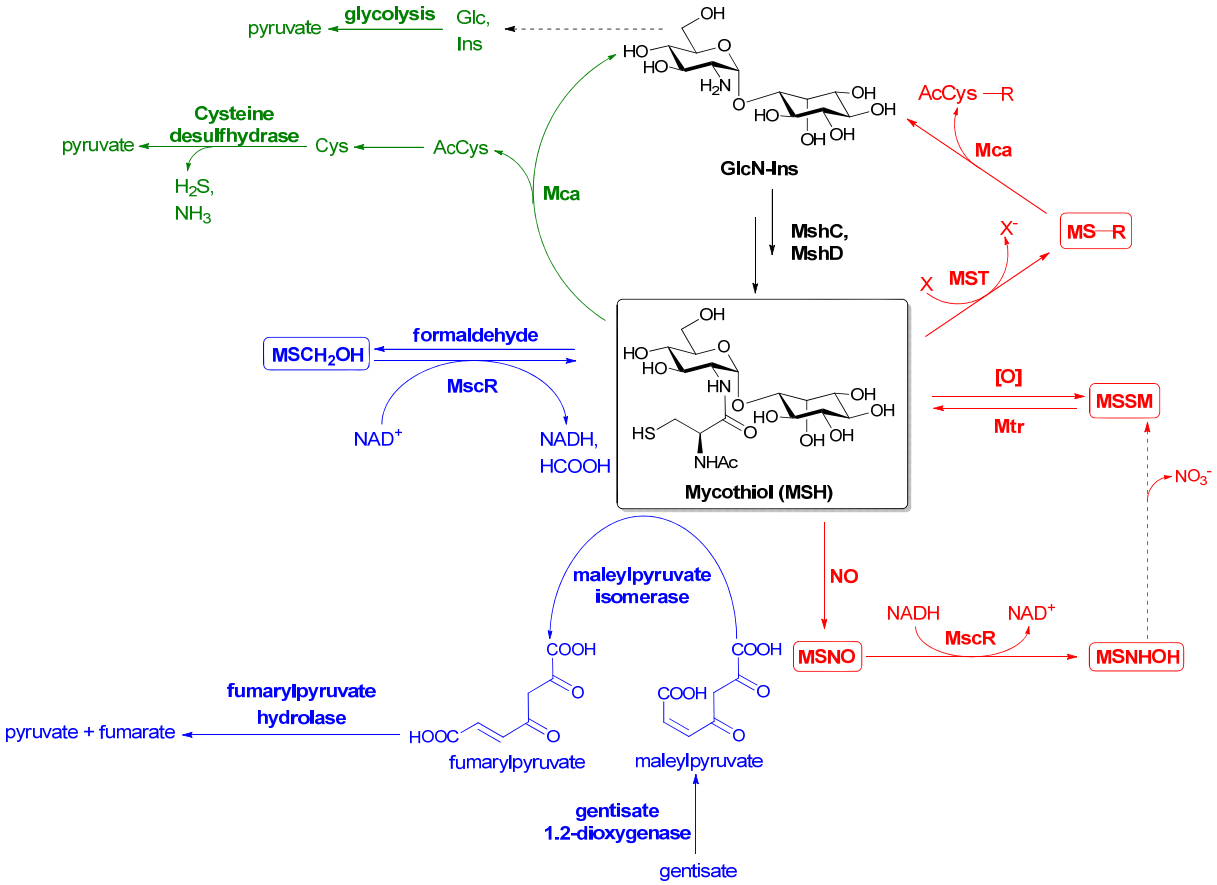


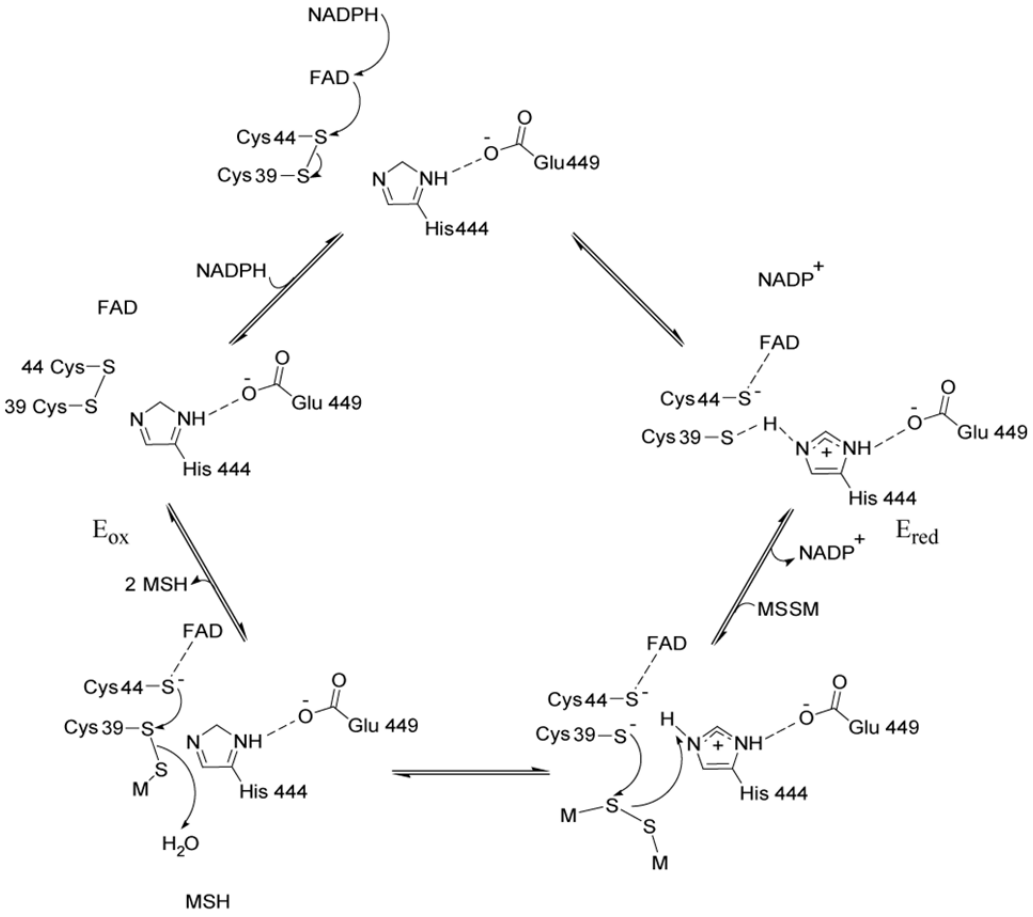
Fig. 1.3: Metabolism reactions of MSH [35]. The metabolic processes of MSH include defensive reactions against ROS, RNS and xenobiotic agents, shown in red; degradation reactions during nutrient starvation, shown in green, and metabolic reactions wherein MSH act as a cofactor for the growth of mycobacteria, shown in blue.

1.3.2.1 Maintenance of the intracellular reducing environment

Mycobacteria utilize MSH to reduce both ROS and RNS [5, 14]. Thiol peroxidases catalyse the reduction of ROS into MSSM and water. MscR reduce MSNO (MSNO is

generated from the reaction of MSH with nitric oxide) to mycothiol-*N*-hydroxysulfenamides, which is ultimately transformed into MSSM and nitrate (Figure 1.3). In addition, MscR assists with the detoxification of formaldehyde. The elimination of the reactive species generates large amounts of MSSM; however, the ratio of MSH to MSSM is crucial to the mycobacteria to maintain their intracellular environment. Mtr, a NADPH-dependent flavoprotein, helps to maintain the MSH/MSSM ratio by reducing MSSM to two MSH molecules.

In 2001, Patel and Blanchard suggested a catalytic mechanism of *M. tuberculosis* Mtr based on pH and kinetic isotope effect studies done with various substrates. This mechanism involves three catalytically active sites situated at His^{444}, Cys^{39} and Cys^{44} (Scheme 1.5) [36]. The binding of the oxidized Mtr to NADPH results in the



Scheme 1.5: The proposed catalytic mechanism of Mtr. Scheme adapted from Patel *et al*., 2001[36].

oxidation of NADPH to $NADP^+$ and the concomitant reduction of Mtr due to the transfer of two electrons to the active-site of the enzyme. His^{444} is thought to stabilize the reduced active-site disulfide of the enzyme. In addition, it shares a common

electron with Cys^{39} which reacts with the substrate, MSSM, resulting in the release of one molecule of MSH. Release of the first MSH is followed by the release of the second MSH which results in the generation of an enzymatic disulfide (oxidized Mtr), completing the mechanism of Mtr. The reductive half reaction is the rate-limiting step in the redox pathway.

1.3.2.2 Detoxification of electrophilic xenobiotics

Detoxification of xenobiotics in actinomycetes is facilitated by MSH and Mca (Figure 13) [37]. This was discovered by Newton *et. al*. [37] in an attempt to alkyate the free thiol of MSH by adding monobromobimane (mBBr) to *M. smegmatis* cultures. Addition of mBBr to the cultures converted MSH into a bimane derivative (MsmB) which was subsequently cleaved into GlcN-Ins and the bimane derivate of *N*-acetylcysteine (AcCySmB). AcCySmB was excreted from the cell into the medium, while GlcN-Ins was recycled in mycothiol biosynthesis. Mca was identified as the enzyme responsible for liberation of GlcN-Ins and AcCySmB, and was characterised as a zinc-dependent *N*-acyl hydrolase with about 42% sequence homology with MshB [14, 37].

The catalytic mechanism of Mca involves the binding of MSH to an electrophilic xenobiotic *via* a glucosaminyl-amide bond, resulting in the formation of MS-conjugates [14, 37]. The conjugate is subsequently hydrolysed at the glucosaminyl-amide bond by Mca, yielding a mercapturic acid derivate and GlcN-Ins. Finally, the mercapturic acid derivative is excreted, while GlcN-Ins is used to reform MSH, through the action of MshC and MshD. Mca cleaves a number of mycothiol conjugates including MSmB, MS-acetophenone and MS-rifamycin through binding to the mycothiol moieties of these compounds. MSmB has been reported to be the best substrate for Mca identified to date. In addition Mca is believed to be the major player in the development of drug resistance in *M. tuberculosis* [14].

1.3.2.3 Mycothiol as a cofactor, cysteine reservoir and/or carbon source

During nutrient starvation mycobacteria utilise gentisate as an alternative carbon source via the gentisate pathway which is catalyse by gentisate 1,2-dioxygenase to give maleylpyruvate which is converted into fumarylpyruvate which is eventually hydrolysed into pyruvate and fumarate by fumarylpruvate hydrolase (Figure 3). The

conversion of maleylpyruvate into fumarylpyruvate is catalysed by maleylpyruvate isomerase, which uses MSH as a cofactor [14, 35, 38, 39]. Furthermore, MSH can act as a reservoir for Cys and GlcNAc. This is achieved when Mca breaks the MSH down into Cys and GlcNAc, to generate pyruvate as a carbon source during the Krebs cycle [14].

1.4 The MSH biosynthetic pathway and MSH-dependent enzymes as potential targets for the development of new anti-TB drugs

Gene knockout studies of the MSH biosynthetic pathway enzymes and MSH-dependent enzymes have shown that MSH is essential for the growth and survival of *M. tuberculosis* [14, 23, 40, 41]. These observations suggest that the MSH biosynthetic enzymes and MSH-dependent enzymes could be potential targets for development of new anti-TB drugs against the latent and active forms of TB. In this regard the MSH-dependent enzyme Mca and the MSH biosynthetic pathway enzyme MshB have specifically been highlighted as potential drug targets [23, 42-45].

1.4.1 Mca and MshB as potential targets

Mca and MshB play an important role in the biosynthesis pathway and utilization of MSH [14, 37]. MshB is responsible for the conversion of GlcNAc-Ins to GlcN-Ins during the biosynthesis pathway, while Mca is responsible for the cleavage of the amide bond of MSH-*S*-conjugates to give GlcN-Ins and AcCys-*S*-conjugates during detoxification process of alkylating agents and toxins material in the mycobacteria system. The GlcN-Ins produced from both the enzymatic reactions is used by MshC and MshD to generate MSH which is then used in several metabolic reactions (Figure 1.3) [14, 35]. MshB and Mca have been highlighted as potential drug target due to several factors. First, Mca and MshB show a high level of similarity in regards to both primary sequence (homology) and mechanism, suggesting that inhibitors developed against the one protein would also target the other. This was in fact found to be the case in recent studies [23, 43]. There is also very little homology between these enzymes and their eukaryotic counterparts, suggesting that specificity of inhibition should be achievable. Mca is also one of the earliest MSH-dependent enzymes to be discovered and assayed, so that much information about it is know,

while the catalytic mechanism and structure of MshB is well known, which is essential for rational drug design [24, 27, 28, 44]. Taken together, these factors have led to various studies that focused on developing inhibitors against these enzymes.

1.4.2 Previous studies focussing on Mca and MshB inhibitors development

Mca has received most attention as a potential drug target since it is one of the earliest MSH-dependent enzymes identified, and based on its uniqueness compared to other known eukaryotic enzymes [41]. The first inhibitor compounds against Mca were isolated from an Australian non-verongid sponge of *Oceanipia* species by the Bewley group in 2001 [41]. They isolated and studied the structure of five novel alkaloids compounds from the marine extract and evaluated their potential as inhibitor compounds against Mca using fluorescence-detected HPLC. Out of the five novel compounds, four were found to show inhibition activity against Mca. This was followed by several other studies wherein derivative compound based on these natural products were synthesized and tested as potential inhibitors of Mca [45, 46]. However, the development of more potent inhibitors of Mca were stalled by the lack of appropriate activity assays and the lack of structural and mechanistic information of the enzyme, which is essential for rational drug design. Subsequently, the purification and characterization of MshB by Fahey group [47] paved the way for the Bewley group to develop the first inhibitor compounds against MshB through molecular docking studies of the active site, wherein the inhibitor compounds were found to display inhibition activity not only on MshB, but also on Mca [43]. Since the publication of these results, subsequent studies have focussed more on the development of drugs against MshB due to the knowledge of its catalytic mechanism and the availability of its structure [42].

1.4.3 Shortcomings and current needs

Currently, all drug development studies against MshB are stalled by poor availability of it natural substrate (GlcNAc-Ins), and the lack of suitable alternative substrates that can be used to assay its activity. Also, the production of inhibitor compounds and substrate analogues by synthetic elaboration of the GlcNAc-Ins scaffold yields only small quantities of material, and even these are difficult to generate. In addition, there

is currently no continuous assay available by which the MshB reaction can be assayed, making it difficult to screen and characterize libraries of inhibitor compounds that target this enzyme. Therefore, to improve the success of studies that target inhibition of MshB (and/or Mca) for the development of new anti-tuberculosis drugs, these shortcomings must be addressed. This can be done by devising new strategies that will enable the large (gram quantity) scale production of GlcNAc-Ins and compounds based on the GlcNAc-Ins scaffold (to increase the availability of the reaction substrate). Additionally, the development of a continuous MshB activity will greatly facilitate the ability to libraries of inhibitory compounds.

1.5 Objectives and aims of the project

This project aimed to address these shortcomings through two main objectives. The **first objective** was to develop a high-throughput continuous assay for MshB that can be used to identify new inhibitors of the enzyme by screening existing compound libraries. The **second objective** was to prepare and characterize a biocatalyst that can be used for the preparation of new alternative substrates and natural substrate of MshB.

The specific aims of the project can therefore be stated as follows:

Aim 1: *Investigation and characterization of known alternative substrates of MshB for use in the development of a new continuous assay of the enzyme.*

A short introduction to the studies that led to the realization of this aim, and the description of the obtained results, are provided in Chapter 2.

Aim 2: *Engineering of an α-N-acetylglucosaminidase for biocatalytic preparation of thioglycosides as alternative substrate of MshB*

The work done on the generation of such α-thioglycoligase biocatalysts is reported in Chapter 3, and is accompanied by a brief introduction to the use of glycosidase enzymes as biocatalysts.

Taken together, the work presented in this thesis represents a significant step forward in our attempt to the MSH biosynthetic enzymes and MSH-dependent enzymes as potential targets for antituberculosis drug development.

1.6 References

1. Flynn JL, Chan J: Immunology of tuberculsosis. *Annual Review of Immunology.* 2001, **19**(1):93-129.
2. Koul A, Arnoult E, Lounis N, Guillemont J, Andries K: The challenge of new drug discovery for tuberculosis. *Nature.* 2011, **469**(7331):483-490.
3. Bhave DP, Muse WB, 3rd, Carroll KS: Drug targets in mycobacterial sulfur metabolism. *Infectious Disorders Drug Targets.* 2007, **7**(2):140-158.
4. Smith I: Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clinical Microbiology Reviews.* 2003, **16**(3):463-496.
5. Lamprecht DA: Development of a drug discovery protocol through the expression of key mycothiol biosynthetic enzymes from *Mycobacterium tuberculosis*. Stellenbosch: University of Stellenosch; 2008,pp. 1-28
6. Schafer FQ, Buettner GR: Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biology and Medicine.* 2001, **30**(11):1191-1212.
7. Trachootham D, Lu W, Ogasawara MA, Nilsa RD, Huang P: Redox regulation of cell survival. *Antioxidants and Redox Signaling.* 2008, **10**(8):1343-1374.
8. Clugston M, Flemming R, Vogt D: Chemistry, 1st Ed. Cape Town: Oxford University Press Southern Africa; 2002.
9. Ahsan MK, Lekli I, Ray D, Yodoi J, Das DK: Redox regulation of cell survival by the thioredoxin superfamily: an implication of redox gene therapy in the heart. *Antioxidants and Redox Signaling.* 2009, **11**(11):2741-2758.
10. Schwander S, Dheda K: Human lung immunity against *Mycobacterium tuberculosis*. *American Journal of Respiratory and Critical Care Medicine.* 2011, **183**(6):696-707.
11. Nakamura H, Nakamura K, Yodoi J: Redox regulation of cellular activation. *Annual Review of Immunology.* 1997, **15**(1):351-369.
12. Storz G, Imlay JA: Oxidative stress. *Current Opinion Microbiology.* 1999, **2**(2):188-194.
13. Sen CK: Redox signaling and the emerging therapeutic potential of thiol antioxidants. *Biochemical Pharmacology.* 1998, **55**(11):1747-1758.
14. Jothivasan VK, Hamilton CJ: Mycothiol: synthesis, biosynthesis and biological functions of the major low molecular weight thiol in actinomycetes. *Natural Product Reports.* 2008, **25**(6):1091-1117.

15. Rawat M, Av-Gay Y: Mycothiol-dependent proteins in actinomycetes. *FEMS Microbiology Review.* 2007, **31**(3):278-292.
16. Newton GL, Fahey RC, Cohen G, Aharonowitz Y: Low-molecular-weight thiols in streptomycetes and their potential role as antioxidants. *Journal of Bacteriology.* 1993, **175**(9):2734-2742.
17. Dosanjh NS, Rawat M, Chung J-H, Av-Gay Y: Thiol specific oxidative stress response in Mycobacteria. *FEMS Microbiology Letters* 2005, **249**(1):87-94.
18. Fan F, Vetting MW, Frantom PA, Blanchard JS: Structures and mechanisms of the mycothiol biosynthetic enzymes. *Current Opinion in Chemical Biology.* 2009, **13**(4):451-459.
19. Bzymek KP, Newton GL, Ta P, Fahey RC: Mycothiol Import by *Mycobacterium smegmatis* and function as a resource for metabolic precursors and energy production. *Journal of Bacteriology.* 2007, **189**(19):6796-6805.
20. Buchmeier NA, Newton GL, Fahey RC: A mycothiol synthase mutant of *Mycobacterium tuberculosis* has an altered thiol-disulfide content and limited tolerance to stress. *Journal of Bacteriology.* 2006, **188**(17):6245-6252.
21. Newton GL, Ta P, Bzymek KP, Fahey RC: Biochemistry of the initial steps of mycothiol biosynthesis. *Journal of Biology Chemistry* 2006, **281**(45):33910-33920.
22. Vetting MW, Frantom PA, Blanchard JS: Structural and enzymatic analysis of MshA from *Corynebacterium glutamicum*. *Journal of Biological Chemistry.* 2008, **283**(23):15834-15844.
23. Newton GL, Buchmeier N, Fahey RC: Biosynthesis and functions of mycothiol, the unique protective thiol of actinobacteria. *Microbiolog and Molecular Biology Reviews.* 2008, **72**(3):471-494.
24. Huang X, Hernick M: Examination of mechanism of *N*-acetyl-1-d-myo-inosityl-2-amino-2-deoxy-α-d-glucopyranoside deacetylase (MshB) reveals unexpected role for dynamic tyrosine. *Journal of Biological Chemistry.* 2012, **287**(13):10424-10434.
25. Newton GL, Av-gay Y, Fahey RC: N-Acetyl-1-d-myo-Inosityl-2-Amino-2-Deoxy-α-D-glucopyranoside deacetylase (MshB) Is a key enzyme in mycothiol biosynthesis. *Journal of Bacteriology.* 2000, **182**(24):6958-6963.
26. Rawat M, Kovacevic S, Billman-Jacobe H, Av-Gay Y: Inactivation of *mshB*, a key gene in the mycothiol biosynthesis pathway in *Mycobacterium smegmatis*. *Microbiology.* 2003, **149**(5):1341-1349.

27. Maynes JT, Garen C, Cherney MM, Newton G, Arad D, Av-Gay Y, Fahey RC, James MNG: The Crystal Structure of 1-D-myo-Inosityl 2-Acetamido-2-deoxy-α-D-glucopyranoside geacetylase (MshB) from *Mycobacterium tuberculosis* reveals a zinc hydrolase with a lactate dehydrogenase fold. *Journal of Biological Chemistry.* 2003, **278**(47):47166-47170.
28. McCarthy AA, Peterson NA, Knijff R, Baker EN: Crystal structure of MshB from *Mycobacterium tuberculosis*, a deacetylase involved in mycothiol biosynthesis. *Journal of Molecular Biology.* 2004, **335**(4):1131-1141.
29. Sareen D, Steffek M, Newton GL, Fahey RC: ATP-dependent l-cysteine:1d-myo-inosityl 2-amino-2-deoxy-α-d-glucopyranoside ligase, mycothiol biosynthesis enzyme MshC, is related to class I cysteinyl-tRNA synthetases. *Biochemistry.* 2002, **41**(22):6885-6890.
30. Fan, Luxenburger A, Painter GF, Blanchard JS: Steady-state and pre-steady-state kinetic analysis of *Mycobacterium smegmatis* cysteine ligase (MshC). *Biochemistry.* 2007, **46**(40):11421-11429.
31. Tremblay LW, Fan F, Vetting MW, Blanchard JS: The 1.6 Å crystal structure of *Mycobacterium smegmatis* MshC: The penultimate enzyme in the mycothiol biosynthetic pathway. *Biochemistry.* 2008, **47**(50):13326-13335.
32. Newton G, Fahey R: Mycothiol biochemistry. *Archives of Microbiology.* 2002, **178**(6):388-394.
33. Vetting MW, Roderick SL, Yu M, Blanchard JS: Crystal structure of mycothiol synthase (Rv0819) from *Mycobacterium tuberculosis* shows structural homology to the GNAT family of N-acetyltransferases. *Protein Science.* 2003, **12**(9):1954-1959.
34. Vetting MW, Yu M, Rendle PM, Blanchard JS: The substrate-induced conformational change of *Mycobacterium tuberculosis* mycothiol synthase. *Journal of Biological Chemistry.* 2006, **281**(5):2795-2802.
35. Lamprechet DA: Studies in mycothiol biosynthesis: Identification, characterization and reconstitution of the *M. tuberculosis* pathway enzymes. Stellenbosch: University of Stellenbosch; 2012.
36. Patel MP, Blanchard JS: Mycobacterium tuberculosis mycothione reductase: pH dependence of the kinetic parameters and kinetic isotope effects. *Biochemistry.* 2001, **40**(17):5119-5126.

37. Newton GL, Av-Gay Y, Fahey RC: A novel mycothiol-dependent detoxification pathway in mycobacteria involving mycothiol S-conjugate amidase. *Biochemistry.* 2000, **39**(35):10739-10746.

38. Feng J, Che Y, Milse J, Yin Y-J, Liu L, Rückert C, Shen X-H, Qi S-W, Kalinowski J, Liu S-J: The gene *ncgl2918* encodes a novel maleylpyruvate isomerase that needs mycothiol as cofactor and links mycothiol biosynthesis and gentisate assimilation in corynebacterium glutamicum. *Journal of Biological Chemistry.* 2006, **281**(16):10778-10785.

39. Shen X-H, Zhou N-Y, Liu S-J: Degradation and assimilation of aromatic compounds by *Corynebacterium glutamicum*: another potential for applications for this bacterium. *Applied Microbiology and Biotechnology.* 2012, **95**(1):77-89.

40. Sareen D, Newton GL, Fahey RC, Buchmeier NA: Mycothiol is essential for growth of *Mycobacterium tuberculosis* Erdman. *Journal of Bacteriology.* 2003, **185**(22):6736-6740.

41. Buchmeier N, Fahey RC: The *mshA* gene encoding the glycosyltransferase of mycothiol biosynthesis is essential in *Mycobacterium tuberculosis* Erdman. *FEMS Microbiology Letters.* 2006, **264**(1):74-79.

42. Gammon DW, Steenkamp DJ, Mavumengwana V, Marakalala MJ, Mudzunga TT, Hunter R, Munyololo M: Conjugates of plumbagin and phenyl-2-amino-1-thioglucoside inhibit MshB, a deacetylase involved in the biosynthesis of mycothiol. *Bioorganic and Medicinal Chemistry.* 2010, **18**(7):2501-2514.

43. Metaferia BB, Fetterolf BJ, Shazad-ul-Hussan S, Moravec M, Smith JA, Ray S, Gutierrez-Lugo M, Bewley CA: Synthesis of natural product-inspired inhibitors of *Mycobacterium tuberculosis* mycothiol-associated enzymes: The first inhibitors of GlcNAc-Ins deacetylase. *Journal of Medicinal Chemistry.*2007, **50(**25): 6326-6336.

44. Nicholas GM, Newton GL, Fahey RC, Bewley CA: Novel bromotyrosine alkaloids: Inhibitors of mycothiol-*S*-conjugate amidase. 2001, *Organic Letters* **3**(10): 1543-1545

45 Nicholas G M, Eckman LL, Newton GL, Fahey RC, Ray S,Bewley CA: Inhibition and kinetics of *Mycobacterium tuberculosis* and *mycobacterium smegmatis* mycothiol-*S*-conjugate amidase by natural product inhibitors. *Bioorganic and Medicinal Chemistry*. 2003, **11** (4): 601-608.

46. Nicholas GM, Eckman LL, Ray S, Hughes RO, Pfefferkorn JA, Barluenga S, Nicolaou KC, Bewley CA: Bromotyrosine-derived natural and synthetic products as inhibitors of mycothiol-S-conjugate amidase. *Bioorganic and Medicinal Chemistry Letters*. 2002, **12**(17): 2487-2490.

47. Newton GL, Ko M, Ta P, Av-Gay Y, Fahey RC: Purification and characterization of *Mycobacterium tuberculosis* 1D-myo-inosityl-2-acetamido-2-deoxy-α-D-glucopyranoside deacetylase, MshB, a mycothiol biosynthetic enzyme. *Protein Expression and Purification.* 2006, **47**(2):542-550.

Chapter 2: Development of a novel high-throughput continuous MshB assay

2.1. Introduction

MshB is a metal-containing GlcNAc-Ins *N*-deacetylase that catalyzes the third step of MSH biosynthesis (Figure 2.1, and Section 1.3.1.2), and which has been highlighted as a potential target for drug development [1]. However, two practical factors have prevented such drug development studies that target MshB from advancing: first, the lack of availability of its substrate, GlcNAc-Ins, and second, the lack of a general continuous assay method that can be used to determine the enzyme's activity. This is partly because the substrates and reaction products of MshB are spectrophotometrically inactive, which makes it difficult to follow the reaction directly. This chapter described the results of the studies performed to develop a general continuous assay method for MshB.

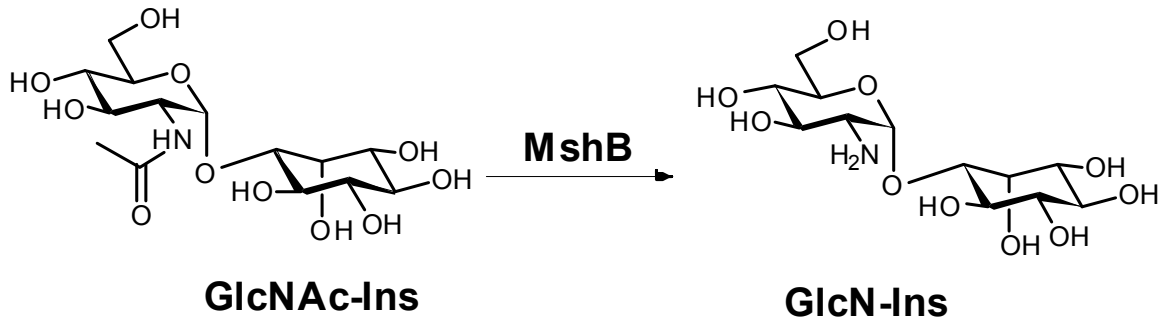


Fig. 2.1: The MshB-catalyzed reaction, in which an acetyl group is removed from GlcNAc-Ins.

To date two discontinuous MshB assays have been developed, both of which are based on the derivatization and subsequent detection of the MshB reaction products. These assays either use the fluorescent derivatization agent 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (AccQ-Fluor™) [2] or fluorescamine (FSA) [3] (Figure 2.2), both of which selectively react with free amine groups. The deacetylation of GlcNAc-Ins by MshB to form GlcN-Ins exposes a free amine group which allows the binding of these fluorescent compounds.

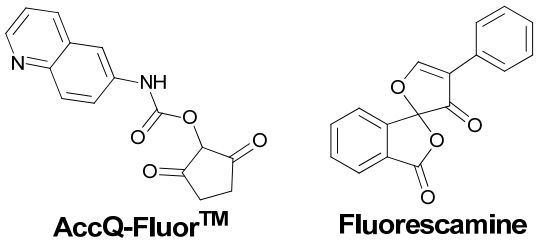


Fig. 2.2: The fluorescent derivatization agents, AccQ-Fluor and fluorescamine, used to analyse the reaction products of MshB.

2.1.1 Previous MSH-related studies based on the use of AccQ-Fluor

AccQ-Fluor was first used by Anderberg *et al*, [2] as a derivatization agent to evaluate the potential metabolites in the MSH biosynthesis pathway of *Mycobacterium smegmatis*. They employed AccQ-Fluor to determine the amounts of glucosamine (GlcN) and GlcN-Ins in cell extracts. The binding of AccQ-Fluor to the amine group of these compounds releases *N*-hydroxysuccinimide (Figure 2.3), after which the reaction mixture is analysed with high-performance liquid chromatography (HPLC).

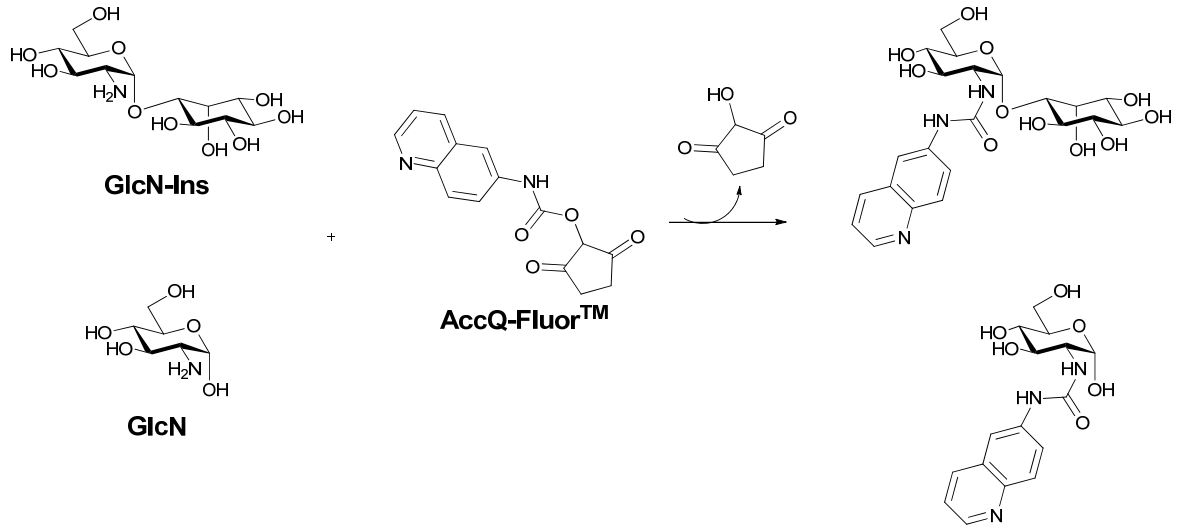
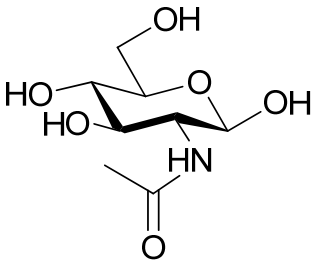


Fig. 2.3: Reaction of AccQ-Fluor with the free amine of GlcN and GlcN-Ins

The AccQ-Fluor™ method was also used by Newton *et al.* [4] to confirm GlcNAc-Ins as a substrate for MshB [4] and for the full characterization of *M. tuberculosis* MshB with GlcNAc-Ins as natural substrate. It was also used to determine the reactivity of MshB with various alternative substrates, including *N*-acetylglucosamine (GlcNAc), a

simplified analogue of GlcNAc-Ins in which the inositol portion of the molecule has been removed (Figure 2.4) [5].



GlcNAc

Fig. 2.4: The structure of *N*-acetylglucosamine

2.1.2 Previous MSH-related studies based on the use of fluorescamine (FSA)

The second derivitization agent, FSA, is a spiro compound that only reacts with primary amines to form a fluorescent product which is detected at λ_{ex} = 390 nm and λ_{em} = 475-490 nm. In 2011 Huang and Hernick [3] used derivatization with FSA to develop an MshB assay in which GlcNAc was used as an alternative substrate (Figure 2.5). For this assay, they followed the deacetylation of GlcNAc over a period of an hour by quenching the reaction at 10 minutes intervals, followed by FSA-derivatization and comparison of the signal to that of a standard curve prepared using GlcN. The obtained data was subsequently used to calculate the enzyme's steady-state parameters.

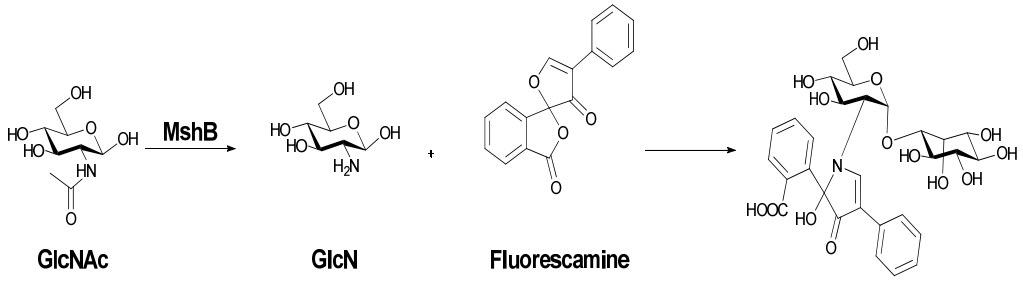


Fig. 2.5: MshB reaction with GlcNAc as substrate, and analysis of its reaction product GlcN by an FSA derivatization-based assay.

2.2 Development of a high-throughput continuous assay of MshB activity

Neither of the assays described above allows the characterization of the MshB activity in a continuous fashion, which limits the ability to screen for new drugs against this enzyme. In addition, discontinuous assays are laborious and make it a cumbersome task to determine the kinetic parameters accuratetly. A previous study by our group identified three chromogenic GlcNAc derivatives, namely 2,4-dinitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose (GlcNAc-SDNP), 4-nitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose (GlcNAc-SPNP) and benzyl 2-acetamido-2-deoxy-1-thio-3,4,6-tri-*O*-acetyl-α-D-glucopyranose (GlcNAc-SBn) as potential alternative substrate for MshB based on HPLC analysis of the reaction mixtures (Figure 2.6) [6]. These compounds differ from the natural MshB substrate GlcNAc-Ins in that the inositol portion was replaced by aromatic groups with absorbance spectra in the UV/Vis-range of the spectrum; they were prepared with the idea that a change in their absorbance spectra would occur upon their deacetylation by MshB, thus allowing the reaction to be followed in a continuous fashion.

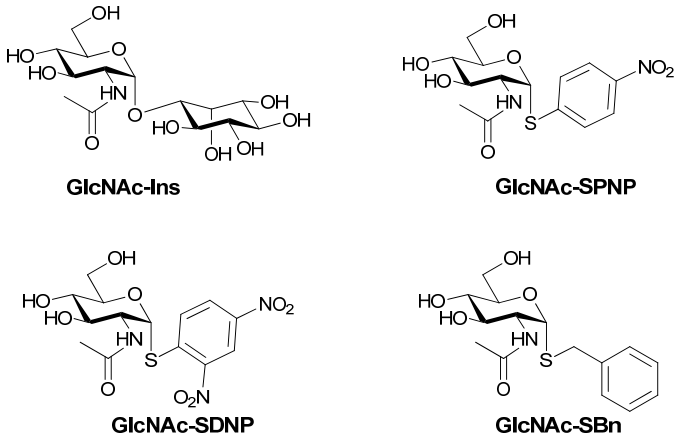


Fig. 2.6: The structure of the MshB natural substrate GlcNAc-Ins, and the three chromogenic GlcNAc derivatives prepared as potential alternative MshB substrates

Unfortunately no significant changes in the absorbance spectra of these compounds were detected during deacetylation by MshB. However, the investigation led to the discovery of an S→N intramolecular rearrangement that occurred upon deacetylation

of one of the compounds, GlcNAc-SDNP. This rearrangement led to the formation of a free SH-group that could react with the known thiol-derivatization agent 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB, or Ellman's reagent). In this manner the alternative substrate could be exploited for the development of the first continuous assay of MshB.

To achieve Aim 1 of this study as described in the previous chapter, GlcNAc-SDNP was characterized as an alternative substrate of MshB using FSA-and DTNB-based activity analyses. Subsequently, a novel, continuous DTNB-based MshB assay was developed based on the S→N intramolecular rearrangement of GlcN-SDNP. The full description of this work, and the results that were achieved was published in May 2012 in *Organic & Biomolecular Chemistry* (*OBC*) [7].

As the second co-author of this publication I was responsible for obtaining the data and results presented in Figure 3A (absorbance spectra), Table 1 (FSA-and DTNB-based activity data) and Figure 5 (activity profiles of GlcNAc-SDNP with MshB using continuous and end-point analysis by means of DTNB derivitazation) of the publication. I also assisted in conducting the experiments done to confirm the S→N intramolecular rearrangement of GlcN-SDNP (MS results presented in Figure 4). The first author, Dr D.A. Lamprecht, was responsible for the chemical synthesis of the chromogenic compounds, as well as the characterization of these compounds as alternative substrate for MshB by means of HPLC analysis (Figure 3B and data in Table 1). Dr. Lamprecht also conducted the experiments done to confirm the S→N intramolecular rearrangement of GlcN-SDNP (LC-ESI-MS data presented in Figure 4), although this was done with my assistance.

This study was done in collaboration with the research group of Dr A. Jardine (UCT) and with the assistance Dr H. Eastwood and Prof. K.J. Naidoo who performed the computational and MshB docking studies (Figure 2). The study was done under the supervision and guidance of Prof. E. Strauss (Dept. of Biochemistry, SU) and Dr. A. Jardine (Dept. of Chemistry, UCT).

The full published version of the *OBC* publication is reproduced on the following pages of this thesis[1]. The supplementary information that accompanied the article is provided as Addendum A.

[1] [DOI: 10.1039/C2OB25429h] – Reproduced by permission of The Royal Society of Chemistry. http://pubs.rsc.org/en/content/articlelanding/2012/ob/c2ob25429h

Organic & Biomolecular Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 5278

www.rsc.org/obc

PAPER

An enzyme-initiated Smiles rearrangement enables the development of an assay of MshB, the GlcNAc-Ins deacetylase of mycothiol biosynthesis†

Dirk A. Lamprecht,[a] Ndivhuwo O. Muneri,[a] Hayden Eastwood,[b,c] Kevin J. Naidoo,[b,c] Erick Strauss*[a] and Anwar Jardine*[c]

Received 27th February 2012, Accepted 17th May 2012
DOI: 10.1039/c2ob25429h

MshB is the *N*-acetyl-1-D-*myo*-inosityl-2-amino-2-deoxy-D-glucopyranoside (GlcNAc-Ins) deacetylase active as one of the enzymes involved in the biosynthesis of mycothiol (MSH), a protective low molecular weight thiol present only in *Mycobacterium tuberculosis* and other actinomycetes. In this study, structural analogues of GlcNAc-Ins in which the inosityl moiety is replaced by a chromophore were synthesized and evaluated as alternate substrates of MshB, with the goal of identifying a compound that would be useful in high-throughput assays of the enzyme. In an unexpected and surprising finding one of the GlcNAc-Ins analogues is shown to undergo a Smiles rearrangement upon MshB-mediated deacetylation, uncovering a free thiol group. We demonstrate that this chemistry can be exploited for the development of the first continuous assay of MshB activity based on the detection of thiol formation by DTNB (Ellman's reagent); such an assay should be ideally suited for the identification of MshB inhibitors by means of high-throughput screens in microplates.

Introduction

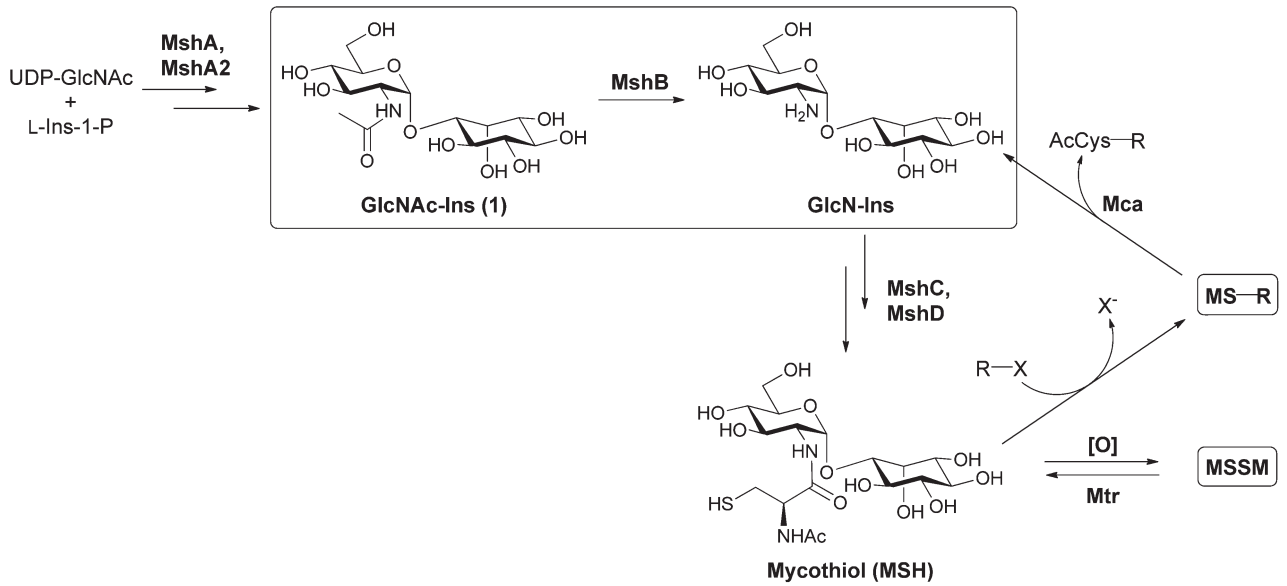
The ability to counteract the effects and consequences of oxidative stress is an important adaptation that ensured the survival of organisms in an aerobic environment.[1] Moreover, many pathogenic microorganisms also maintain virulence by resisting the oxidative killing mechanisms of the human immune system's defences. Low molecular weight (LMW) thiols play a key role in these processes by acting as redox buffers, thereby helping to maintain redox homeostasis within cells.[2] In most organisms—including humans—the principle LMW thiol is glutathione. However, many Gram-positive bacteria do not contain glutathione, and evidence suggests that they instead rely on the metabolic cofactor coenzyme A (CoA) and/or the recently discovered bacillithiol (BSH) as redox buffer.[3] The actinomycetes, which include *Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis, produce mycothiol (MSH) as their principal LMW thiol.[4,5] These differences, and the essential requirement for MSH under various stress conditions, led to the identification of the biosynthesis and utilization of this novel thiol as a potential target for tuberculosis drug development.[6]

MSH is biosynthesised from UDP-GlcNAc, L-*myo*-inositol-1-phosphate (L-Ins-1-P), L-cysteine and acetyl-CoA through the sequential action of four enzymes, designated MshA, MshB, MshC and MshD (Scheme 1).[4,7] A fifth enzyme, an as yet unidentified phosphatase tentatively named MshA2, has also been implicated in the reaction.[8] Two other enzymes also play important roles in the maintenance of MSH levels: the NADPH-dependent oxidoreductase mycothiol disulfide reductase (Mtr) is responsible for the reduction of mycothiol disulfide (MSSM) formed through the oxidation of MSH,[9,10] while the mycothiol-*S*-conjugate amidase (Mca) cleaves the cysteine amide of mycothiol-*S*-conjugates formed in the MSH-dependent detoxification of alkylating agents, free radicals and xenobiotics. In the process it forms 1-D-*myo*-inosityl-2-amino-2-deoxy-D-glucopyranoside (GlcN-Ins) and a mercapturic acid derivative, which is excreted.[11]

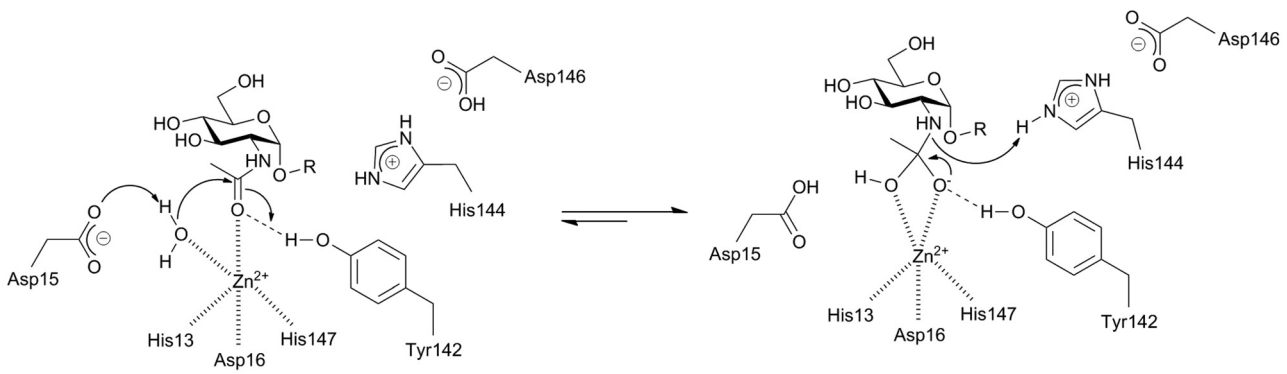
The mycothiol biosynthetic enzyme MshB is one of the enzymes in the pathway that has been identified as an attractive potential target for drug discovery. This was based on the fact that MSH levels in knockout mutants of the mycobacterium *M. smegmatis* that lacked the *mshB* gene were reduced up to 95% compared to wild-type (the low residual formation of MSH was ascribed to the promiscuous nature of Mca).[12] MshB is a metal-dependent hydrolase[13] that catalyses the deacetylation of *N*-acetyl-1-D-*myo*-inosityl-2-amino-2-deoxy-D-glucopyranoside (GlcNAc-Ins, **1**) to form the amino sugar GlcN-Ins. Structural

[a]*Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa. E-mail: estrauss@sun.ac.za; Fax: +2721 808 5863; Tel: +27 21 808 5866*
[b]*Scientific Computing Research Unit, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa*
[c]*Department of Chemistry, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa. E-mail: anwar.jardine@uct.ac.za; Fax: +2721 650 4010; Tel: +27 21 650 4010*

†Electronic supplementary information (ESI) available: Full chromatograms and associated mass spectra from the LC-HR-ESI-MS analyses, and NMR spectra of the thioglycosides **3a–c**. See DOI: 10.1039/c2ob25429h



Scheme 1 Biosynthesis and recycling of mycothiol (MSH).



Scheme 2 Proposed catalytic mechanism of MshB, highlighting the active site residues involved in catalysis.[17]

analysis showed that the active site metal—previously thought to be zinc, although recent studies have found higher activity with iron[14]—is coordinated to two histidine residues (His13 and His147) and an aspartate (Asp16).[15,16] The metal is also bound to the acetamidocarbonyl oxygen of the substrate and a water molecule, which was proposed to be activated for nucleophilic attack through this interaction. This has been confirmed in recent mechanistic studies, which also showed that MshB uses a general acid–base mechanism in which Asp15 and His144 act as the general base and acid catalysts respectively, while Tyr142 plays a dynamic role that includes stabilization of the oxyanion intermediate (Scheme 2).[17]

The study of MshB activity and inhibition poses two major practical challenges. First, its natural substrate can only be obtained by means of lengthy multi-step synthetic procedures,[18,19] or by isolation of mycothiol-*S*-conjugates from natural sources, followed by treatment with Mca to release GlcN-Ins (Scheme 1).[20] Consequently, many recent MshB studies made use of commercially available GlcNAc (**2**) as an alternate minimal substrate instead,[14,17,21] even though it has been shown that MshB is ~100-fold less active towards this compound.[4,20] Additionally, other groups have invested significant synthetic effort in the discovery of alternate MshB substrates that give reasonable levels of activity.[22] Second, the general method for the assay of MshB activity involves derivatization of reaction products followed by lengthy HPLC analysis, which precludes high-throughput screening.[22–24] A significant advancement has been the recent development of a fluorescence-based microplate assay that measures MshB activity based on the reaction of fluorescamine (FSA) with the free amine of the deacetylation product after the reaction has been quenched.[21] However, its mechanism precludes its use for the screening of compound libraries for potential MshB inhibitors in which similarly reactive functional groups are represented.

In this study we sought to address these shortcomings by the incorporation of chromogenic functional groups into the structures of potential alternate MshB substrates. Our strategy was to exploit any changes that occur in the absorbance spectra of such compounds upon deacetylation for use in a continuous spectrophotometric assay. In the process we uncovered an unexpected enzyme-initiated Smiles rearrangement of one of the substrates, which allowed the development of a new, sensitive microplate-based assay of MshB activity.

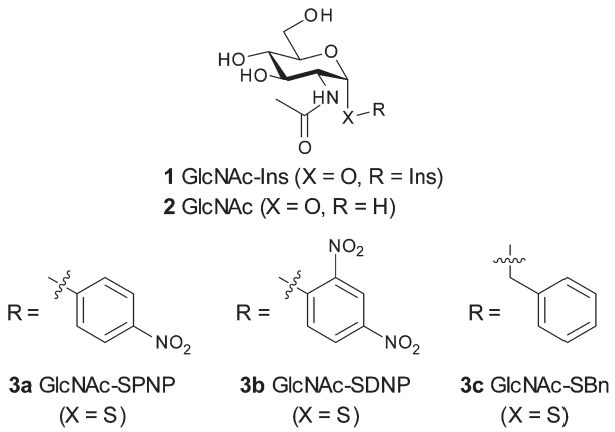


Fig. 1 Structures of alternate MshB substrates used in this study.

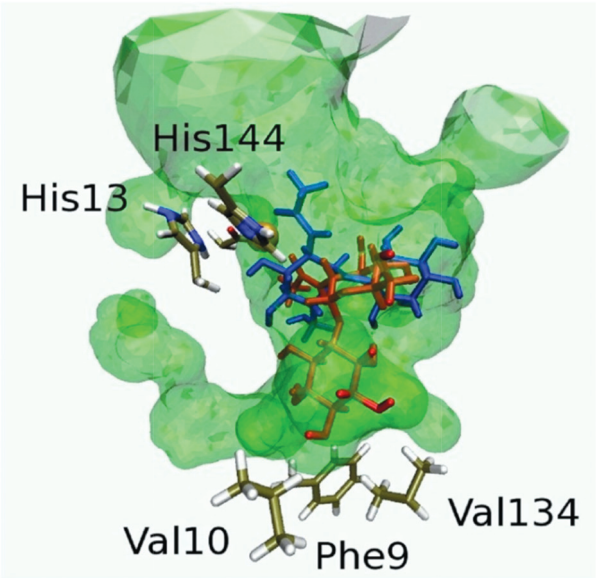


Fig. 2 Molecular docking analysis of the MshB active site binding pocket (enclosure shown in green) identifies two possible binding orientations for the native substrate GlcNAc-Ins (**1**), shown in red (geometry 1) and blue (geometry 2) respectively. His13 which coordinates zinc, and the general acid catalyst His144 are shown, as well as the hydrophobic residues at the base of the pocket.

Results and discussion

Rationale and strategy

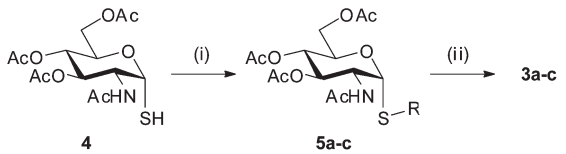
Previous studies of MshB and Mca have shown that modification of the inosityl moiety of GlcNAc-Ins is usually tolerated by these enzymes, although often to the detriment of the observed activity.[19,22] We therefore chose to investigate whether the replacement of the inosityl group by aromatic moieties would similarly result in compounds accepted as substrates by MshB. More specifically, we set out to use aromatic groups that are known strong chromophores, anticipating that deacetylation of such alternate substrates may result in changes in their absorbance maxima. Such changes would allow the continuous spectrophotometric assaying of MshB activity, which would be ideal for high-throughput screening. We decided to introduce such groups *via* *S*-glycoside linkages, since the syntheses of several α-GlcNAc thioconjugates have been reported.[19,22,24–26]

Three aromatic groups were chosen as replacements of the inosityl group of GlcNAc-Ins for this study (Fig. 1). In the first case a *para*-nitrophenyl group, a well-known chromophore utilized in several different enzymatic assays, was selected to give GlcNAc-SPNP (**3a**). Second, a 2,4-dinitrophenyl group was chosen to give GlcNAc-SDNP (**3b**) as a potential alternate substrate. The extra nitro-group was expected to increase the absorbance intensity of the chromophore, thereby potentially also increasing its sensitivity to any changes that might occur in its absorbance spectrum upon deacetylation. Third, a simple benzyl group was selected as a reference for the study of the other two groups, as its absorbance was unlikely to be affected by deacetylation due to its distant location from the acetamido group. Moreover, a recent study found that GlcNAc-SPh (the phenyl thioglycoside) showed ~30% of the activity of the native GlcNAc-Ins substrate, suggesting that GlcNAc-SBn (**3c**) might similarly act as a good alternate substrate, thereby also acting as a positive control.[22]

Docking studies and computational chemistry

One of the alternate substrates that we proposed to use in this study, the dinitrophenyl derivative **3b**, was recently reported to be inactive in a MshB assay that depended on the fluorescent derivatization of the free amine of the expected product, followed by HPLC analysis.[22] While this result indicated that **3b** would not be suitable for the purposes of this study, it seemed curious in light of the good activity observed for the structurally analogous GlcNAc-SPh using the same assay. Evaluation of the available structures of recombinant MshB indicated that its binding pocket is cavernous and that it should be able to accommodate a variety of substrates, as indicated by the green enclosure shown in Fig. 2.[15,16] However, charged or polar amino acids predominate in the binding cavity, with few well-defined hydrophobic regions. This suggested that the hydrophobic aromatic groups of the alternate substrates would not be accommodated in the MshB active site, although this could be offset by the potential of the nitro substituents to act as hydrogen bond acceptors. To establish whether **3a** and/or **3b** would be excluded from MshB's active site based on such structural considerations, we set out to perform docking and molecular dynamics studies using one of the available MshB structures (PDB: 1Q74).[16] While the active site zinc ion is clearly visible in this structure, it does not contain any other bound ligands. The natural substrate GlcNAc-Ins (**1**) and minimal substrate GlcNAc (**2**) were therefore also included in the analysis for comparative purposes.

The results of the molecular dynamics studies performed with the solution-equilibrated MshB structure indicated that the docked MshB–ligand complexes with good scores have the acetamidocarbonyl oxygen in close proximity to the zinc, which is consistent with the current mechanistic understanding of metal-dependent deacetylases.[13] Moreover, the relatively large active site cavity volume permits two favourable docked-ligand geometries for each of the ligands, both of which have the acetamidocarbonyl oxygen placed in a manner allowing catalysis to take place. For the natural substrate GlcNAc-Ins, the two binding



Scheme 3 Synthesis of thioglycosides **3a–c**. (i) 1-fluoro-4-nitrobenzene (for **a**), 1-fluoro-2,4-dinitrobenzene (for **b**) or benzyl bromide (for **c**), Et_3N, CH_2Cl_2, 25 °C. (ii) MeOH, H_2O, acetone, Amberlite IRA-400 (OH^-). The R-groups are defined in Fig. 1.

pocket orientations can be defined using its coordination to the zinc ion as pivot, and its displacement relative to His13 and Asp16 as reference points (Fig. 2). In geometry 1, GlcNAc-Ins extends into a pocket facing down and away from His13, while in geometry 2 GlcNAc-Ins stretches across into a binding pocket that is orthogonal to geometry 1. The results indicate that for GlcNAc-Ins there may be a small but significant energetic preference for geometry 1 over geometry 2.

In regards to the alternate substrates, the docking analysis indicates that the minimal substrate GlcNAc (**2**) prefers binding in geometry 2, which may partially explain its reduced activity. Similarly GlcNAc-SPNP (**3a**) also prefers binding in geometry 2, while GlcNAc-SDNP (**3b**) prefers geometry 1. The analysis also shows that regardless of the binding geometry, the aromatic rings of both proposed alternate substrates can be accommodated in the same pocket that binds inositol. However, the presence of the hydrophobic residues Val10, Phe9 and Val134 at the base of the pocket that surrounds geometry 1 would seem to indicate that this would be the binding mode most likely to compensate for the loss of the various hydrogen bonding interactions when the inosityl moiety of GlcNAc-Ins in replaced by an aromatic group.

Taken together, the results of the docking studies suggest that **3a** and **3b** would most likely act as alternate substrates for MshB, with activities similar to that previously observed for GlcNAc-SPh.[22]

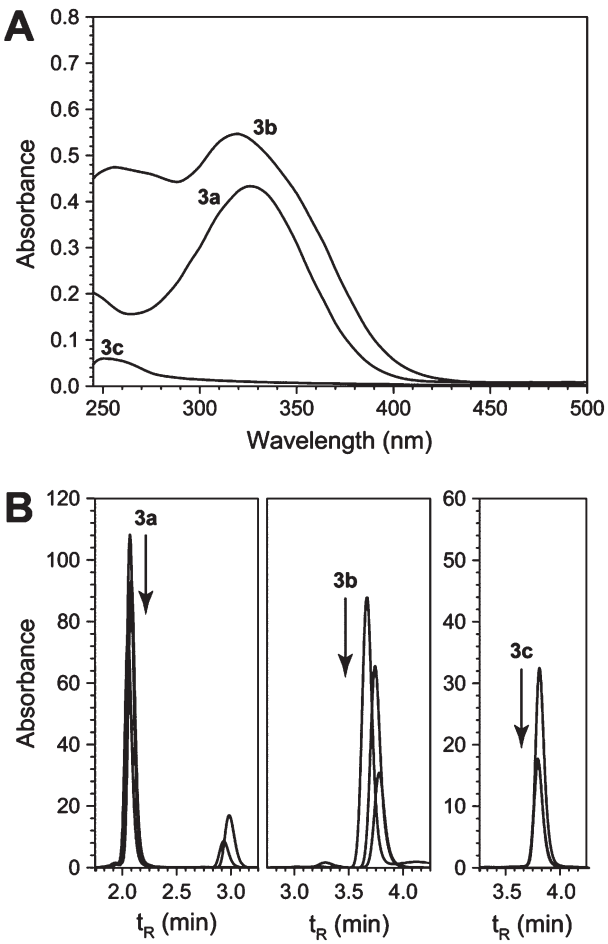


Fig. 3 Characterization of **3a–c** as potential alternate MshB substrates. Panel A, UV-vis spectra of **3a–c** at 200 μM. Panel B, HPLC analysis of 100 μM each of **3a**, **3b** (at 320 nm) and **3c** (at 254 nm) incubated in the presence of MshB followed by heat inactivation at set time intervals, shows a time-dependent decrease in the concentration of each compound.

Synthesis of potential chromogenic substrates of MshB

The α-GlcNAc-mercaptan (**4**) has previously been used in the synthesis of a range of α-thioglycosides—including analogues of MSH—by simple nucleophilic substitution using an appropriate electrophile.[19,26] We therefore decided to expand on this established chemistry for the synthesis of the α-thioglycosides **3a–c** (Scheme 3). Triethylamine-mediated *S*-alkylation of **4** with aryl fluorides or benzyl bromide gave the desired *O*-acetylated thioglycosides (**5a–c**). Subsequent *O*-deacetylation gave the unprotected thioglycosides (**3a–c**). As expected, the coupling constants of the anomeric protons were all consistent with that of α-coupled thioglycosides.

Confirming MshB activity with alternate substrates by HPLC

With the thioglycosides **3a–c** in hand we set out to demonstrate that these are indeed accepted by MshB as alternate substrates. The UV-visible spectra of **3a–c** showed that, as predicted, the nitrophenyl-containing **3a** and **3b** have large extinction coefficients at ~325 nm and ~320 nm respectively, while the benzyl-containing **3c** has a relatively poor absorbance with a maximum at ~254 nm (Fig. 3A). Nonetheless, the presence of the chromophoric groups in all three cases allowed their direct analysis by HPLC, which provided for a simple method whereby their potential deacetylation by MshB could be monitored. Subsequent incubation of MshB individually with 100 μM each of **3a–c** showed a time-dependent decrease for all three compounds, confirming that they indeed act as alternate substrates of the enzyme (Fig. 3B). Integration of the peak area at each time point allowed for the HPLC results to be converted to time-course data by comparison to a standard curve prepared independently. In this manner the specific activity of the enzyme towards each substrate could be calculated (Table 1). These values indicated that MshB has a similar specific activity for all three the alternate substrates, which is ~12% of the value previously determined for the natural substrate GlcNAc-Ins using HPLC analysis of the fluorescent derivative of the product.[27] However, it is about an order of magnitude better than the value determined for the minimal substrate GlcNAc in the same manner, indicating that

Table 1 MshB activity of various substrates as measured by using the indicated assays[a]

Assay	Substrate	Specific activity (nmol min^{-1} mg^{-1})	Rate (V/E)[b] (min^{-1})
HPLC (AccQ-Fluor deriv.)[c]	GlcNAc-Ins (**1**)	220 ± 40[27]	7.21 ± 1.31[27]
	GlcNAc (**2**)	3 ± 2[27]	0.10 ± 0.07[27]
HPLC (direct)[d]	GlcNAc-SPNP (**3a**)	26.9 ± 1.3	0.88 ± 0.04
	GlcNAc-SDNP (**3b**)	28.4 ± 3.0	0.93 ± 0.10
	GlcNAc-SBn (**3c**)	25.4 ± 3.0	0.83 ± 0.10
Fluorescamine (FSA)[e]	GlcNAc (**2**)	19.1 ± 3.4	0.63 ± 0.11
	GlcNAc-SPNP (**3a**)	~0	~0
	GlcNAc-SDNP (**3b**)	4.2 ± 3.7[f]	0.14 ± 0.12
	GlcNAc-SBn (**3c**)	122.2 ± 17.0[f]	4.01 ± 0.56
DTNB[e]	GlcNAc (**2**)	~0	~0
	GlcNAc-SPNP (**3a**)	~0	~0
	GlcNAc-SDNP (**3b**)	145.8 ± 15.9	4.78 ± 0.52
	GlcNAc-SBn (**3c**)	~0	~0

[a] Activity values determined in this study represent the average of three independent measurements, with the standard error shown. [b] Calculated from the specific activity data. [c] Using 100 μM substrate in water, and 0.5 μM MshB. [d] Using 100 μM substrate in water, and 2 μM MshB. [e] Using 5 mM substrate in 5–10% DMSO, and 1 μM MshB. [f] Based on a standard curve prepared with GlcN.

the thioglycosides **3a–c** may act as suitable alternative substrates for the assay of MshB activity.

Confirmation of the formation of the expected products upon deacetylation of 3a–c by MshB

Since synthetic standards of the expected deacetylation products of **3a–c** were not available, the HPLC-based MshB activity assay described above could only be used to confirm the disappearance of substrate and not the formation of these products. To confirm that MshB treatment of **3a–c** yields products containing free amino groups, we employed the recently reported MshB assay in which its substrate (GlcNAc in the original description) is incubated in the presence of the reactive fluorophore FSA.[21] The FSA reacts with the free amine of the deacetylated product as it is formed to give a fluorescent conjugate, thereby providing for the monitoring of product formation—and consequently MshB activity—by fluorimetry.

The activity of MshB towards GlcNAc and the thioglycosides **3a–c** (5 mM in 10% DMSO) were monitored in this manner over a period of 1 h. The observed increase in fluorescence for each substrate over time was subsequently converted to a change in concentration over time using a standard curve prepared from known concentrations of glucosamine (GlcN); in doing so, the assumption was made that the reactivity of the deacetylated products of **3a–c** towards fluorescamine would be similar to GlcN, and that the rates obtained in this manner (after adjustment for any background using a control sample containing no substrate) would allow for a direct comparison of the activity of MshB towards the respective substrates.

Surprisingly, the resulting data (Table 1) gave rise to a very different activity profile than the one determined by HPLC analysis. The FSA-based assay highlights GlcNAc-SBn **3c** as being the preferred MshB substrate among the thioglycoside, showing a specific activity ~6-fold that of GlcNAc **2**, while GlcNAc-SPNP **3a** and GlcNAc-SDNP **3b** show no or very low activities (with large errors). Since the HPLC-based assay clearly indicated the disappearance of all three thioglycosides, this finding suggested that the products formed upon deacetylation of **3a** and **3b** do not contain free amine groups that are available for reaction with FSA. We next set out to establish a possible mechanistic basis for such an outcome.

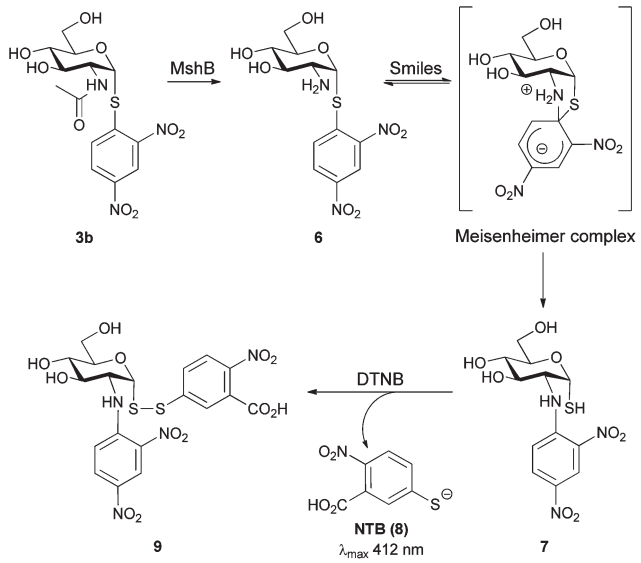
Deacetylation of GlcNAc-SDNP unmasks a free thiol: an enzyme-initiated Smiles rearrangement

We considered that the unmasking of the free amino group of **3a** and **3b** could lead to a possible intramolecular rearrangement that results in this group being blocked again, thereby preventing their reaction with FSA. Such a case was reported by Kondo *et al.*, which found that *S*-(2,4-dinitrophenyl)cysteine undergoes a facile base-catalyzed S → N rearrangement to form *N*-(2,4-dinitrophenyl) cysteine.[28] This type of rearrangement, which falls under a broad category of intramolecular aromatic nucleophilic substitution reactions, involves the formation of an anionic σ complex intermediate, also referred to as a Meisenheimer complex. Such rearrangements were first reported by Smiles in 1930, and are therefore known by his name.[29] In the experiment by Kondo *et al.*, the reaction was performed by incubating *S*-(2,4-dinitrophenyl)cysteine with different organic bases, including imidazole. This suggested that enzyme-borne bases could also promote such Smiles rearrangements, and that this could be the basis for the surprise finding that **3a** and **3b** are seemingly inactive in the FSA-based MshB assay.

A Smiles rearrangement of the thioglycoside substrates after deacetylation by MshB should unmask a free thiol group, as shown for **3b** in Scheme 4. Since the concentration of free thiols can be established through reaction with 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB, or Ellman's reagent), GlcNAc (acting as negative control) and **3a–c** was individually incubated with MshB. The reactions were stopped at set time intervals, followed by treatment with DTNB and measurement of the increase in absorption at 412 nm to determine the rate of formation of 2-nitro-5-thiobenzoate (NTB, **8**), if any.

The results (Table 1) showed an increase only in the case of GlcNAc-SDNP **3b**, with none of the other substrates showing an increase above the background rate (no substrate control). Moreover, the rate observed for **3b** in this assay is very similar to the rate seen for **3c** in the FSA-based assay. This finding confirms that GlcNAc-SDNP **3b** does act as an alternate substrate for MshB, and that its apparent inactivity (based on the results of the FSA-based assay and as reported in a previous study that used AccQ-Fluor derivatization followed by HPLC analysis) is due to an S → N Smiles rearrangement.

The observation that GlcNAc-SPNP **3a** does not show a similar unmasking of its thiol, while also being unreactive in the



Scheme 4 Proposed Smiles rearrangement of the MshB-catalysed deacetylation product (**6**) of GlcNAc-SDNP **3b**. The formation of the rearrangement product **7** can be followed by derivatisation with DTNB (Ellman's reagent) to form the strong chromophore, 2-nitro-5-thiobenzoate (NTB, **8**) and the mixed disulfide **9**.

FSA assay, may be due to the reversible formation of the corresponding Meisenheimer complex that is less stabilized by the single nitro group, and which therefore does not complete the rearrangement. Since the docking studies indicated that **3a** and **3b** may prefer to take on different binding modes, this might also play a role in differentiating their activity profiles. However, no direct evidence in support of either analysis can be provided at this stage.

Identification of the Smiles rearrangement product

We subsequently set out to positively identify and study the Smiles rearrangement product by its chemical synthesis; however, all attempts in this regard failed, as all the conditions used for *N*- and/or *O*-deacetylation led to the formation of *N*-(2,4-dinitrophenyl) glucosamine and H_2S. We therefore attempted its identification by LC-HR-ESI-MS analysis instead. Comparison of reaction mixtures that contained **3b** in either the absence or presence of MshB indicated the enzyme-mediated formation of the deacetylation product **6** but not its rearrangement product **7** (Fig. 4). This identification is based on the analysis of the associated mass spectrum, which showed the formation of fragments that is in agreement with the structure of **6** but not **7**, although it is also possible that a stabilized Meisenheimer complex would give the same observed fragmentation pattern (see ESI†); note that all three structures (**6**, **7**, and the complex) have the same molecular mass. Addition of the reductant tris-(2-carboxyethyl)phosphine (TCEP) to the reaction mixtures gave the same result. To establish whether the rearrangement product is formed, but decomposes prior to or during analysis, DTNB was added to mixtures after they had been quenched and the precipitated protein removed. Samples that had been treated in this

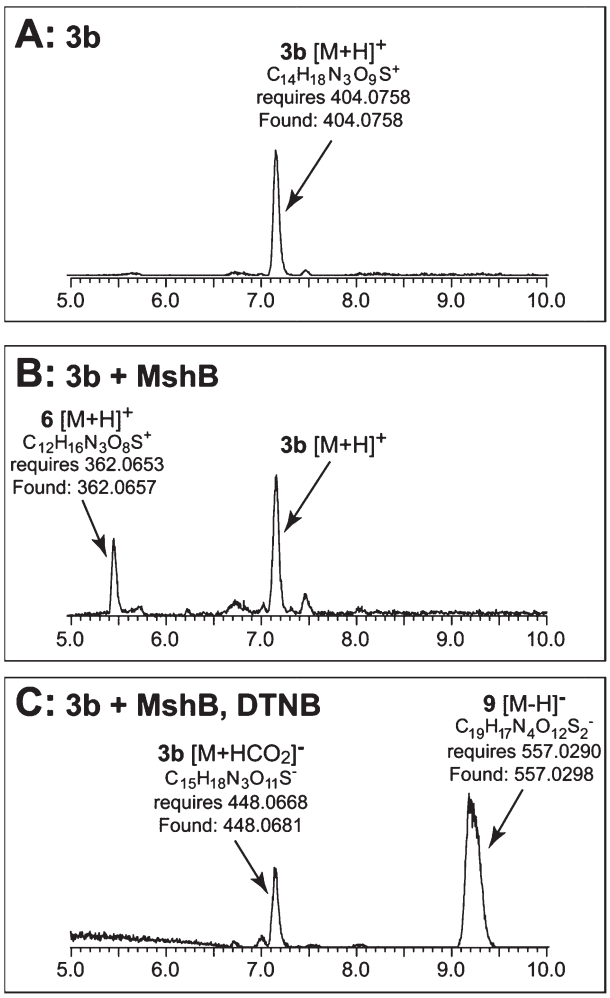


Fig. 4 LC-HR-ESI-MS analyses of the product(s) formed in the MshB-mediated reaction of GlcNAc-SDNP **3b**. Panel A, base peak ion (BPI) chromatogram of a control reaction mixture containing only the substrate **3b** in the presence of the reductant TCEP. Panel B, BPI chromatogram of a reaction mixture containing **3b**, MshB and TCEP. Panel C, BPI chromatogram of a reaction mixture containing **3b** and MshB (but no TCEP) that had been treated with DTNB prior to analysis. All peaks were identified by analysis of the associated mass spectra (the required and found mass for the major ion in each spectrum is shown) and observed fragmentation patterns. The full chromatograms and the associated mass spectra are provided as ESI.†

manner, and that did not contain TCEP, showed the formation of a new peak that was identified as the mixed disulfide **9** (Fig. 4); in the presence of TCEP this peak was not observed.

The results of the LC-HR-ESI-MS analyses confirm the deacetylation of GlcNAc-SDNP **3b** by MshB to form the deacetylation product **6**, as well as its subsequent rearrangement to form **7**. Since the FSA-based assay failed to trap any free amine-containing products, it is most likely that the observation of **6** in these analyses is due to its formation from the stabilized Meisenheimer complex under the acidic conditions used for the MS analysis; the Meisenheimer complex could also give the same fragmentation pattern directly (Fig. 4B). Formation of the labile rearranged product **7** is confirmed by its trapping by DTNB to

form the stable disulfide adduct **9**; however, in the absence of a trapping reagent **7** is not observed, in agreement with the difficulties encountered with its chemical synthesis.

Exploiting the Smiles rearrangement to develop a new MshB assay amenable to high-throughput screening

The confirmed deacetylation of GlcNAc-SDNP **3b** by MshB led us to consider whether this substrate could be used for the development of a new assay for MshB activity, which is also amenable to high-throughput screening. Towards this end, we first determined whether product formation could be measured based on changes that occur in the absorbance spectrum of **3b** upon its deacetylation (and/or subsequent Smiles rearrangement). However, incubation of **3b** in the presence of MshB showed no significant time-dependent changes in its absorbance at 320 nm (its absorbance maximum, see Fig. 3A), making direct spectrophotometric-based activity analysis impossible and necessitating an alternative strategy.

The reaction of DTNB with free thiols have been used to assay the activity of several enzymes (including the mycothiol disulfide reductase, Mtr[9]) in continuous fashion. Having confirmed the reaction of DTNB with **7** and the formation of the mixed disulfide **9** and the associated NTB (**8**) by-product, we therefore set out to determine if MshB activity could also be determined in this manner. Increasing concentrations of **3b** (between 0 and 5 mM) were incubated in the presence of MshB and DTNB, and the increase in absorbance at 412 nm monitored in a time-dependent fashion. For comparison, identical reaction mixtures were incubated without DTNB, stopped at regular time intervals, and the amount of product formed determined by DTNB treatment after the enzyme was precipitated and removed as before.

The progress curves obtained from the continuous monitoring of MshB activity through NTB release showed an initial lag period (presumably due to the slow collapse of the Meisenheimer complex) before giving way to the linear increase in NTB formation. Importantly, a control reaction that contained 100 μM substrate but no enzyme showed no increase, confirming that the observed increase in absorbance was due to enzyme activity. Regression analysis of the linear portions of the progress curves allowed for the corresponding activity rates to be calculated, while the rates from the discontinuous assay were determined from regression analysis of the data obtained at the set time-intervals at each concentration. The results (Fig. 5) show that the data obtained by either method give comparable rates at substrate concentrations of 1 and 5 mM. However, the continuous monitoring of activity is clearly superior since it gave data points with smaller errors, leading to an excellent linear correlation between activity and substrate concentration over the whole concentration range measured. In fact, when the activity of reaction mixtures with less than 1 mM substrate was determined in a discontinuous fashion, the variation in the replicates became so large that accurate measurements became impossible (data not shown). The activity obtained using 5 mM **3b** (3.4 ± 0.1 min^{-1} from the continuous assay, and 5.0 ± 1.0 min^{-1} for the discontinuous assay) also correlate well with those obtained for 5 mM GlcNAc–SBn **3c** obtained in the FSA-based assay (Table 1). This suggests that

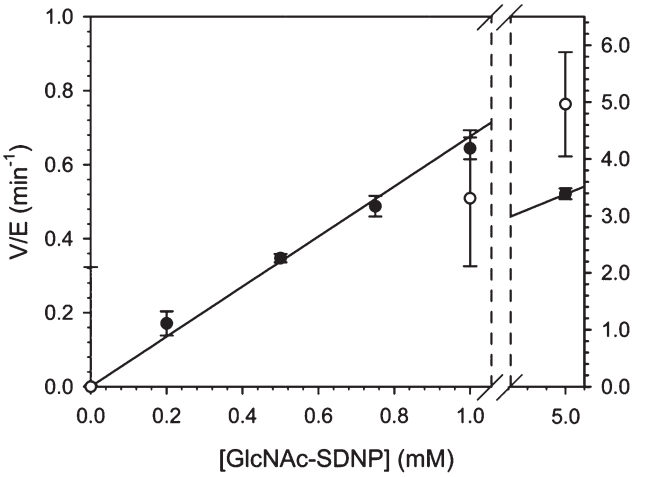


Fig. 5 Activity profile of MshB with GlcNAc-SDNP **3b** as substrate, measured based on the release of NTB (**8**) after reaction of DTNB (Ellman's reagent) with the rearranged deacetylation product **7** (see Scheme 4). Closed circles (●) represent data points determined through the continuous monitoring of product formation by performing the reaction in the presence of DTNB; Open circles (○) are data points obtained by quenching reaction mixtures are regular intervals, followed by DTNB treatment. All data points represent the average of two separate measurements, with the errors bars indicating the standard deviation. Linear regression analysis of the (●) data set gives a line with the equation $y = 0.6763x + 0.000138$. The *y*-axis labels on the left and right are associated with the data points to the left and right of the *x*-axis break respectively; the units are identical in both cases.

under the assay conditions used, the kinetics of the Smiles rearrangement do not adversely affect the measurement of MshB's deacetylation of **3b** by DTNB derivatization.

The finding that MshB's activity towards **3b** as a substrate correlates linearly over the concentration range used (0–5 mM) was not surprising considering that the K_M value determined for GlcNAc using the FSA assay is 38 ± 4 mM (k_{cat} 46 ± 2.2 min^{-1}).[21] However, the poor solubility of **3b** in aqueous solutions prevented us from using higher concentrations, and therefore the kinetic parameters could not be determined for this substrate. Nonetheless, the rate obtained for 5 mM GlcNAc-SDNP using the DTNB-based assay is ~7–12% of the k_{cat}(GlcNAc) obtained in the FSA-(46 ± 2.2 min^{-1})[21] and HPLC-based (29.4 ± 2.4 min^{-1})[27] assays respectively, suggesting that the kinetic parameters for **3b** do not differ significantly from those for GlcNAc.

Finally, the result demonstrates that GlcNAc-SDNP **3b** acts as a suitable alternative substrate for MshB, allowing its activity to be measured continuously using a DTNB-based assay at concentrations as low as 50 μM.

Conclusions

In this study we set out to develop a new continuous assay of MshB activity based on the use of chromogenic substrates. In the process, we uncovered an unexpected and previously unseen enzyme-initiated Smiles rearrangement of one of the substrates.

This discovery allowed for the first continuous and sensitive assay of MshB activity to be developed.

The study of the mechanistic detail of the Smiles rearrangement of deacetylated **3b** would be interesting considering not only the novelty of the enzyme-based initiation of the reaction, but would also help relate the rate of deacetylation to the kinetic and thermodynamic factors that control the formation and subsequent collapse of the associated Meisenheimer complex. Moreover, it is also unclear whether the Smiles rearrangement takes place in the MshB active site, where interaction between the unmasked thiol and the catalytic metal ion may result in inhibition, or only after the deacetylated product is released. However, it is quite likely that the rearrangement takes place upon collapse of the tetrahedral intermediate, and that adjacent acidic residues such as His144 plays a role in stabilizing the Meisenheimer complex intermediate. Conversely, the very rigid platform provided by the glucosamine skeleton, and the presence of stabilising electron-withdrawing nitro substituents in **3b**, would promote the rearrangement independent of enzyme involvement. A full understanding of the interplay between such considerations would be helpful in further establishing the utility of the assay.

In conclusion, the serendipitous discovery of GlcNAc-SDNP **3b** as an alternate MshB substrate with unique characteristics provides the opportunity for microplate-based inhibitor screening of this enzyme, or any other deacetylase reactivity that may potentiate the Smiles rearrangement by similarly invoking heteroatom nucleophiles. This finding may significantly advance our search of inhibitors of mycothiol biosynthesis as potential antituberculosis agents.

Materials and methods

General reagents and methods

Chemicals were purchased from Sigma-Aldrich. All reactions were performed at room temperature unless specified otherwise. Flash chromatography was performed using a 20 : 1 weight ratio of silica to crude reaction mixture. Thin layer chromatography (TLC) was performed on Merck Aluminium Silica gel 60 F_{254} plates. All compounds containing free amines were visualized on TLC plates by spraying the plates with ninhydrin solution (1.5% ninhydrin in *n*-butanol with 3% AcOH) followed by heating the plates until the colour developed. All chromogenic compounds were visualized on the TLC plate under UV light. HPLC analyses were performed on an Agilent 1100 system with an in-line DAD detector. LC-HR-ESI-MS analyses were performed using a Waters Aquity UPLC attached to a Waters Synapt G2 QTOF mass spectrometer. High resolution ESI-MS analyses were performed on a QTOF Ultima API quadrupole mass spectrometer. ^{1}H and ^{13}C NMR was performed on Varian Unity Inova 600 MHz NMR and Varian VXR 300 MHz NMR spectrometers. Enzyme assays (fluorometric and spectrophotometric analyses) were conducted using a Varioskan multimode reader (Thermo Scientific). An extinction coefficient (ε) of 13 600 dm^{-3} mol^{-1} cm^{-1} was used for the absorbance of 2-nitro-5-thiobenzoate (NTB, **8**) at 412 nm.

Computational methods

Prior to performing the docking calculations the MshB crystal structure (PDB ID: 1Q74) was protonated using the Karlsberg pK_a prediction tool.[30] The protein was hydrated for molecular dynamics (MD) simulations using a sphere of 30.0 Å in a manner similar to one reported previously.[31]

MD simulations were performed using the CHARMM32b[32] program with the CHARMM27 all atom force field for the protein[33] and CSFF for the carbohydrate and inositol.[34] In all MD simulations a water sphere of 30.0 Å containing 11 322 TIP3P[35] water molecules as implemented in CHARMM[36] were used. The system was centred about the glycosidic linkage of the ligand for the simulations. The water molecules that overlapped with any of the ligand heavy atoms were removed and a spherical boundary force was applied to the surface of the water sphere to maintain a consistent water density of 1.0 g cm^{-1}. A constant temperature of 298.15 K was maintained throughout the simulations using the Nosé–Hover thermostat.[37] A buffer region of 2 Å thickness and a dynamic region with 21 Å radius was used with Langevin dynamics in all simulations. All hydrogen bond lengths were kept fixed with the SHAKE algorithm and an integration step of 1 fs was used.

The crystal structure was subjected to 7 ns explicitly solvated molecular dynamics and then stripped of water molecules, after which several ligands were docked in the MD solution equilibrated MshB structure using Glide.[38] A cubic volume of side 9 Å, centred on the central zinc atom was selected. A search was conducted *via* a grid generation with no constraints in conjunction with a Van der Waals scaling factor of 1 and a partial charge cut-off of 0.25. Conformational searches on each ligand were performed using the *Extra Precision* (XP) algorithm, where a maximum of 2000 conjugate gradient steps, a distance dielectric cut-off of 2 Å and flexible docking were applied on 1200 poses for every ligand.

Of the conformations generated, those where the Coulomb and van der Waals interactions summed to values greater than 0 were rejected. To ensure that the ligands generated were conformationally distinct an RMS deviation of 0.5 Å was used. The 100 highest scoring geometries for each molecule were examined to identify structures where the carbonyl carbon is in close proximity to the zinc.

Synthesis of α-thioglycoside compounds

4-Nitrophenyl 2-acetamido-2-deoxy-1-thio-3,4,6-tri-*O*-acetyl-α-D-glucopyranose (5a). To a solution of compound **4**[26] (280 mg, 0.773 mmol) in 5 ml dichloromethane stirred under a N_2 atmosphere was added 1-fluoro-4-nitrobenzene (164 µl, 2.8 mmol) and triethylamine (130 µl, 1.15 mmol). The reaction was allowed to stir for 24 h, and then concentrated. The product was purified from the crude reaction mixture using flash chromatography (2% MeOH/CH_2Cl_2), resulting in 360 mg (96% yield) of pure **5a** as a yellow amorphous solid. δ_H (600 MHz, $CDCl_3$, Me_4Si): 1.974 (s, 3H), 2.026 (s, 3H), 2.061 (s, 3H), 2.081 (s, 3H), 4.070 (dd, 1H), 4.254 (dd, 1H), 4.373 (m, 1H), 4.590 (m, 1H), 5.169 (m, 2H), 5.902 (d, 1H), 6.013 (d, 1H), 7.560 (d, 2H), 8.152 (d, 2H); δ_C (150 MHz, CD_3COCD_3): 21.9, 23.9, 54.3,

64.0, 71.1, 71.5, 72.3, 87.6, 126.0, 132.0, 145.5, 171.8, 171.9 and 207.4. *m/z* (100% ESI-MS): 485 $[M + H]^+$, 507 $[M + Na]^+$.

2,4-Dinitrophenyl 2-acetamido-2-deoxy-1-thio-3,4,6-tri-*O*-acetyl-α-D-glucopyranose (5b). To a solution of compound **4**[26] (500 mg, 1.37 mmol) in 10 ml dichloromethane stirred under a N_2 atmosphere was added 1-fluoro-2,4-dinitrobenzene (173 µl, 1.37 mmol) and triethylamine (230 µl, 1.65 mmol). The mixture was allowed to stir under a N_2 atmosphere for 24 h. The reaction mixture was then concentrated and the resulting product was purified using flash chromatography (2% $MeOH/CH_2Cl_2$), resulting in 570 mg (79% yield) of pure **5b** as a yellow solid. δ_H (300 MHz, $CDCl_3$, Me_4Si): 1.935 (s, 3H), 1.991 (s, 3H), 2.034 (s, 3H), 2.066 (s, 3H), 4.043 (dd, 1H), 4.265 (m, 2H), 4.626 (m, 2H), 5.235 (m, 2H), 6.100 (d, 1H), 8.045 (d, 1H), 8.406 (dd, 1H), 9.026 (d, 1H); *m/z* (100% ESI-MS): 530 $[M + H]^+$.

Benzyl 2-acetamido-2-deoxy-1-thio-3,4,6-tri-*O*-acetyl-α-D-glucopyranose (5c). To a solution of compound **4**[26] (220 mg, 0.61 mmol) in 2 ml dichloromethane stirred under a N_2 atmosphere was added benzylbromide (87 µl, 0.732 mmol) and triethylamine (170 µl, 1.22 mmol). The reaction mixture was allowed to stir under a N_2 atmosphere for 24 h. Subsequently the reaction was concentrated and the resulting product was purified using flash chromatography (2% $MeOH/CH_2Cl_2$), resulting in 170 mg (62% yield) of pure **5c** as an off white oil. δ_H (400 MHz, CD_3OD, Me_4Si): 1.89 (s, 3H), 1.93 (s, 3H), 2.01 (s, 3H), 2.06 (s, 3H), 3.81 (d, 2H), 3.98 (d, 1H), 4.24 (dd, 1H), 4.38 (m, 3H), 4.97 (dd, 1H), 5.16 (dd, 1H), 5.29 (d, 1H) and 7.30 (m, 5H). δ_C (100 MHz, CD_3OD, Me_4Si): 20.4, 22.1, 35.0, 52.9, 63.1, 69.3, 70.6, 71.9, 83.8, 128.1, 129.4, 129.9, 138.9 and 171.2.

4-Nitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose (3a). To a solution of **5a** (230 mg, 0.475 mmol) in 18 ml of H_2O (19%), MeOH (50%) and acetone (31%), was added 2.3 g of Amberlite IRA-400 (OH^-). The reaction mixture was allowed to stir for 1 h at ambient temperature. The volatile solvents were removed *in vacuo* and the residual aqueous product mixture was subjected to C_{18} flash chromatography (50% $MeCN/H_2O$). Aqueous fractions containing the compound were collected and dried by lyophilization, resulting in 50 mg (29% yield) of pure **3a** as an amorphous yellow solid. δ_H (300 MHz, CD_3OD, Me_4Si): 1.99 (s, 3H), 3.44 (t, 1H), 3.74 (m, 3H), 3.97 (m, 1H), 4.15 (dd, 1H), 6.08 (d, 1H), 7.68 (d, 2H) and 8.15 (d, 2H). δ_C (75 MHz, CD_3OD, Me_4Si): 22.5, 56.1, 62.4, 72.3, 72.6, 75.7, 87.3, 124.9, 130.6, 146.3, 147.5 and 173.9. *m/z* (100% ESI-HRMS): 359.0929 ($[M + H]^+$ $C_{14}H_{19}N_2O_7S^+$ requires 359.0913).

2,4-Dinitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-gluco-pyranose (3b). Prepared as for **3a** from **5b** (230 mg, 0.434 mmol), resulting in 60 mg (34% yield) of **3b** as an amorphous yellow solid. δ_H (300 MHz, $(CD_3)_2SO_2$, Me_4Si): 1.829 (s, 3H), 3.264 (m, 1H), 3.516 (q, 1H), 3.588 (m, 2H), 3.674 (m, 1H), 3.983 (m, 1H), 4.554 (t, 1H), 5.125 (d, 1H), 5.272 (d, 1H), 6.051 (d, 1H), 8.150 (d, 1H), 8.273 (d, 1H), 8.426 (dd, 1H), 8.818 (d, 1H). *m/z* (100% ESI-HRMS): 404.0781 ($[M + H]^+$ $C_{14}H_{18}N_3O_9S^+$ requires 404.0764).

Benzyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose (3c). Prepared as for **3a** from **5c** (230 mg, 0.507 mmol), resulting in 76 mg (38% yield) of **3c**. m.p. 192–194 °C (dec). δ_H (300 MHz, CD_3OD, Me_4Si): 1.89 (s, 3H), 3.36 (dd, 1H), 3.60 (dd, 1H), 3.76 (m, 4H), 3.99 (m, 2H), 5.34 (d, 1H), 7.20 (t, 1H), 7.27 (t, 2H) and 7.33 (d, 2H). δ_C (75 MHz, CD_3OD, Me_4Si): 22.5, 35.0, 55.6, 62.6, 72.6, 74.5, 84.0, 128.0, 129.4, 130.1, 139.6, 173.5 and 173.5. *m/z* (100% ESI-HRMS): 328.1211 ($[M + H]^+$ $C_{15}H_{22}NO_5S^+$ requires 328.1219).

MshB preparation

The *mshB* (Rv1170) gene was amplified from *M. tuberculosis* H37Rv genomic DNA using as forward primer 5′-GGTGCCTCCATGGCTGAGACGCCGC-3′ and as reverse primer 5′-CCTGGTTGGCCTCGAGCGTGCCGGAC-3′. The primers respectively contained NcoI and XhoI restriction enzyme sites (underlined), while the reverse primer also removed the native stop codon. The resulting PCR product was treated with NcoI and XhoI and ligated to similarly treated pET28a(+) to yield the expression vector pET28a(+)-*mshB* that encodes recombinant MshB with a C-terminal 6×His-tag. The plasmid's sequence was confirmed by DNA sequencing.

For the expression of MshB the pET28a(+)-*mshB* plasmid was transformed (heat shock method) into *E. coli* BL21*(DE3) cells. LB medium (500 ml) was inoculated with a 5 ml starter culture of these cells grown in the same medium, followed by cell growth at 37 °C until the culture reached an OD_{600} of 0.6. Expression was induced by the addition of IPTG to a final concentration of 0.10 mM, after which the culture was incubated overnight at 22 °C. Cells were harvested by centrifugation at 17 000 × *g* for 20 min at 10 °C. The resulting cell pellet was resuspended in binding buffer (20 mM Tris–HCl, 5 mM imidazole, 500 mM NaCl, pH 7.9), and the cells were disrupted using sonication with external cooling (ice bath). Cell debris were removed by centrifugation at 75 000 × *g* for 20 min at 10 °C to obtain the clarified cell extract used for purification.

The protein was purified using an automated purification program on an ÄKTA*prime* system (Amersham Bioscience) fitted with a 1 ml HiTrap Chelating HP column (GE Healthcare) pre-loaded with Zn^{2+}. The column-bound MshB was eluted with 100% elution buffer (20 mM Tris-HCl, 500 mM imidazole, 500 mM NaCl, pH 7.9) and desalted using a 5 ml HiTrap desalting column (GE Healthcare) with gel filtration buffer (25 mM Tris-HCl, 5 mM $MgCl_2$, 5% glycerol, pH 8) on the ÄKTA*prime* system. Protein purity was confirmed with SDS-PAGE analyses and protein concentration was determined with a Bradford assay. Pure MshB (in gel filtration buffer with 5% glycerol added) was aliquoted and stored at −80 °C until needed.

HPLC-based activity analysis

All samples used for the direct HPLC analysis of MshB activity contained one of the thioglycoside substrates (**3a**, **3b** or **3c**) at a final concentration of 0.1 mM, assay buffer (50 mM HEPES, 50 mM NaCl, <1% DMSO, pH 7.4) and MshB (16.7 µg; 2 µM) in a total volume of 250 µl. Reactions were initiated with the addition of MshB. Samples of the reaction were stopped at

intervals of 5 min over a period of 20 min by heat inactivation, followed by removal of the denatured enzyme by centrifugation at 12 000 × *g* for 5 min to obtain the clarified samples used for HPLC analysis.

For the determination of the substrate concentrations in the samples, standard solutions of **3a–c** with a concentration range of 0–120 μM and a final volume of 250 μl each were given the same treatment as the enzyme reaction samples before these were also analysed by HPLC. From these results a standard curve which correlated peak area with concentration was obtained for each substrate.

Samples of 5 μl each were injected on to a Synergi 4μ Fusion-RP 250 × 2.00 mm column (Phenomenex) and eluted using aqueous acetonitrile as solvent. The amount of acetonitrile used differed for each substrate: for GlcNAc-SPNP (**3a**), 30% MeCN/H_2O; for GlcNAc-SDNP (**3b**), 25% MeCN/H_2O; and for GlcNAc-SBn (**3c**), 40% MeCN/H_2O. Separation was monitored at 320 nm for **3a** and **3b**, and at 254 nm for **3c**.

FSA- and DTNB-based activity analyses

Reaction mixtures (250 μl) containing the respective substrate (5 mM) in assay buffer (50 mM HEPES, 50 mM NaCl, 1 mM TCEP, 5–10% DMSO, pH 7.5) and MshB (8.1 μg; 1 μM) were incubated at 30 °C. Samples (30 μl) of the reaction mixture were added at 10 min intervals to 10 μl trichloroacetic acid (TCA) (1.22 M) to quench the reaction, followed by removal of the precipitated protein by centrifugation at 12 000 × *g* for 5 min. The supernatant (25 μl) was subsequently added to 75 μl borate buffer (1 M, pH 9).

For the fluorescamine (FSA)-based analysis, 49 μl of the resulting mixture was added to 30 μl of FSA (10 mM) contained in a 96-well microplate incubated at room temperature, after which the fluorescence was determined using an excitation wavelength of 395 nm, and an emission wavelength of 485 nm.

For the 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB)-based analysis, 49 μl of the resulting mixture was diluted to 100 μl, and then added to 30 μl DTNB (2.2 mM) contained in a 96-well microplate. After incubation at room temperature for 10 min, the absorbance at 412 nm was determined.

LC-HR-ESI-MS analyses

Reaction mixtures (100 μl) contained GlcNAc-SDNP (**3b**) (0.7 mM) in assay buffer (50 mM HEPES, 50 mM NaCl, ~5% DMSO, pH 7.5) (the enzyme was omitted from a negative control reaction). One set of reaction mixtures also contained 1 mM TCEP. Mixtures were equilibrated at 30 °C for 10 min before MshB (1 μM final concentration) was added, followed by incubation at 30 °C for 1 h. The reaction was quenched by adding two 45 μl aliquots from each mixture to two separate microfuge tubes containing 15 μl TCA (6.12 M) each. The precipitated protein was removed by centrifugation at 12 000 × *g* for 5 min, after which the supernatants from each tube (37.5 μl) were added to 112.5 μl borate buffer (1 M, pH 9) contained in two separate microfuge tubes. The tubes subsequently received either 45 μl acetonitrile or 45 μl DTNB (0.72 mM in acetonitrile) prior to LC-HR-ESI-MS analysis.

Reaction mixtures were analysed on a Waters Aquity UPLC system with an autosampler using a Synergi 4μ Fusion-RP 250 × 2.00 mm column (Phenomenex). The sample injection volume was 20 μl and the autosampler syringe was washed with solvent A (0.1% formic acid in water) before each injection. A gradient elution program with a flow-rate of 0.30 ml min^{-1} was used for the analysis. The gradient was as follows: 0–6 min, linear increase to 50% solvent B (0.1% formic acid in acetonitrile); 6–8 min, isocratic 50% B; 8–10 min, linear increase to 95% solvent B. Detection was performed on a Waters Synapt G2 QTOF mass spectrometer. Analytes were detected in the positive and negative ion mode using a capillary voltage of 3000 V, cone voltages of 15 V and a source temperature of 150 °C. The mass spectrometer was calibrated with sodium formate and Leu-enkephalin as internal standard to acquire the lock mass for the HR-ESI-MS analysis. Data processing was done using MassLynx (Waters).

Continuous DTNB-based MshB assay

Reaction mixtures (125 μl) contained GlcNAc-SDNP (**3b**) (0–5 mM) in assay buffer (50 mM HEPES, 50 mM NaCl, ~10% DMSO, pH 7.5) with DTNB (0.75 mM) and MshB (10 μg; 2.5 μM) were incubated at 30 °C in a 96-well microplate, and the increase in absorbance at 412 nm was monitored over a period of 30 min. Reaction rates were determined by linear regression of the resulting progress curves obtained between 5–30 min. Discontinuous reaction monitoring was performed as before, using the same reaction mixture without DTNB and withdrawing 30 μl aliquots at set time intervals for processing and analysis as described above.

Acknowledgements

This work was funded by grants from the Bill and Melinda Gates Foundation (Grant#52035GA) and the National Research Foundation (NRF) of South Africa (Grant#65527) to AJ. The computational work is based upon research supported by the South African Research Chairs Initiative (SARChI) of the Department of Science and Technology (DST) and NRF awarded to KJN. DAL and NOM are grateful recipients of Harry Crossley and NRF scarce skills bursaries respectively.

References

1 J. A. Imlay, *Annu. Rev. Microbiol.*, 2003, **57**, 395–418; J. A. Imlay, *Annu. Rev. Biochem.*, 2008, **77**, 755–776.
2 P. Zuber, *Annu. Rev. Microbiol.*, 2009, **63**, 575–597.
3 G. L. Newton, K. Arnold, M. S. Price, C. Sherrill, S. B. Delcardayre, Y. Aharonowitz, G. Cohen, J. Davies, R. C. Fahey and C. Davis, *J. Bacteriol.*, 1996, **178**, 1990–1995; R. C. Fahey, *Annu. Rev. Microbiol.*, 2001, **55**, 333–356; G. L. Newton, M. Rawat, C. J. J. La, V. K. Jothivasan, T. Budiarto, C. J. Hamilton, A. Claiborne, J. D. Helmann and R. C. Fahey, *Nat. Chem. Biol.*, 2009, **5**, 625–627; J. D. Helmann, *Antioxid. Redox Signaling*, 2011, **15**, 123–133.
4 V. K. Jothivasan and C. J. Hamilton, *Nat. Prod. Rep.*, 2008, **25**, 1091–1117.
5 G. L. Newton, N. Buchmeier and R. C. Fahey, *Microbiol. Mol. Biol. Rev.*, 2008, **72**, 471–494.
6 M. Rawat, C. Johnson, V. Cadiz and Y. Av-Gay, *Biochem. Biophys. Res. Commun.*, 2007, **363**, 71–76.

7 C. Bornemann, M. A. Jardine, H. S. C. Spies and D. J. Steenkamp, *Biochem. J.*, 1997, **325**, 623–629.
8 G. L. Newton, P. Ta, K. P. Bzymek and R. C. Fahey, *J. Biol. Chem.*, 2006, **281**, 33910–33920.
9 C. J. Hamilton, R. M. J. Finlay, M. J. G. Stewart and A. Bonner, *Anal. Biochem.*, 2009, **388**, 91–96.
10 C. M. Holsclaw, W. B. Muse, III, K. S. Carroll and J. A. Leary, *Int. J. Mass Spectrom.*, 2011, **305**, 151–156; M. P. Patel and J. S. Blanchard, *Biochemistry*, 2001, **40**, 5119–5126; K. S. E. Ung and Y. Av-Gay, *FEBS Lett.*, 2006, **580**, 2712–2716.
11 M. Steffek, G. L. Newton, Y. Av-Gay and R. C. Fahey, *Biochemistry*, 2003, **42**, 12067–12076.
12 M. Rawat, S. Kovacevic, H. Billman-Jacobe and Y. Av-Gay, *Microbiology*, 2003, **149**, 1341–1349; X. Xu, C. Vilcheze, Y. Av-Gay, A. Gomez-Velasco and W. R. Jacobs, Jr., *Antimicrob. Agents Chemother.*, 2011, **55**, 3133–3139.
13 M. Hernick and C. A. Fierke, *Arch. Biochem. Biophys.*, 2005, **433**, 71–84.
14 X. Huang, E. Kocabas and M. Hernick, *J. Biol. Chem.*, 2011, **286**, 20275–20282.
15 A. A. McCarthy, N. A. Peterson, R. Knijff and E. N. Baker, *J. Mol. Biol.*, 2004, **335**, 1131–1141.
16 J. T. Maynes, C. Garen, M. M. Cherney, G. Newton, D. Arad, Y. Av-Gay, R. C. Fahey and M. N. G. James, *J. Biol. Chem.*, 2003, **278**, 47166–47170.
17 X. Huang and M. Hernick, *J. Biol. Chem.*, 2012, **287**, 10424–10434.
18 G. M. Nicholas, L. L. Eckman, P. Kovac, S. Otero-Quintero and C. A. Bewley, *Bioorg. Med. Chem.*, 2003, **11**, 2641–2647; M. A. Jardine, H. S. C. Spies, C. M. Nkambule, D. W. Gammon and D. J. Steenkamp, *Bioorg. Med. Chem.*, 2002, **10**, 875–881; K. Ajayi, V. V. Thakur, R. C. Lapo and S. Knapp, *Org. Lett.*, 2010, **12**, 2630–2633; S. Lee and J. P. N. Rosazza, *Org. Lett.*, 2004, **6**, 365–368.
19 S. Knapp, S. Gonzalez, D. S. Myers, L. L. Eckman and C. A. Bewley, *Org. Lett.*, 2002, **4**, 4337–4339.
20 G. L. Newton, Y. Av-Gay and R. C. Fahey, *J. Bacteriol.*, 2000, **182**, 6958–6963.
21 X. Huang and M. Hernick, *Anal. Biochem.*, 2011, **414**, 278–281.
22 D. W. Gammon, D. J. Steenkamp, V. Mavumengwana, M. J. Marakalala, T. T. Mudzunga, R. Hunter and M. Munyololo, *Bioorg. Med. Chem.*, 2010, **18**, 2501–2514.
23 G. M. Nicholas, L. L. Eckman, G. L. Newton, R. C. Fahey, S. Ray and C. A. Bewley, *Bioorg. Med. Chem.*, 2003, **11**, 601–608.
24 B. B. Metaferia, B. J. Fetterolf, S. Shazad-ul-Hussan, M. Moravec, J. A. Smith, S. Ray, M.-T. Gutierrez-Lugo and C. A. Bewley, *J. Med. Chem.*, 2007, **50**, 6326–6336.
25 S. Knapp, B. Amorelli, E. Darout, C. C. Ventocilla, L. M. Goldman, R. A. Huhn and E. C. Minnihan, *J. Carbohydr. Chem.*, 2005, **24**, 103–130; B. Paul and W. Korytnyk, *Carbohydr. Res.*, 1984, **126**, 27–43; Y. Ding, Y. Miura, J. R. Etchison, H. H. Freeze and O. Hindsgaul, *J. Carbohydr. Chem.*, 1999, **18**, 471–475.
26 S. Knapp and D. S. Myers, *J. Org. Chem.*, 2002, **67**, 2995–2999; S. Knapp and D. S. Myers, *J. Org. Chem.*, 2001, **66**, 3636–3638.
27 G. L. Newton, M. Ko, P. Ta, Y. Av-Gay and R. C. Fahey, *Protein Expression Purif.*, 2006, **47**, 542–550.
28 H. Kondo, F. Moriuchi and J. Sunamoto, *J. Org. Chem.*, 1981, **46**, 1333–1336.
29 W. E. Truce, E. M. Kreider and W. W. Brand, *Org. React.*, 2011, 99–215.
30 B. Rabenstein and E.-W. Knapp, *Biophys. J.*, 2001, **80**, 1141–1150; G. Kieseritzky and E.-W. Knapp, *Proteins: Struct., Funct., Bioinf.*, 2008, **71**, 1335–1348.
31 K. J. Naidoo, D. Denysyk and J. W. Brady, *Protein Eng., Des. Sel.*, 1997, **10**, 1249–1261; K. J. Naidoo and J. W. Brady, *THEOCHEM*, 1997, **395–396**, 469–475.
32 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217; B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545–1614.
33 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
34 M. Kuttel, J. W. Brady and K. J. Naidoo, *J. Comput. Chem.*, 2002, **23**, 1236–1243.
35 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
36 P. J. Steinbach and B. R. Brooks, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 9135–9139.
37 S. Nosé, *Mol. Phys.*, 1984, **52**, 255–268; W. G. Hoover, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1985, **31**, 1695–1697.
38 B. Hammouda, *J. Polym. Sci., Part B: Polym. Phys.*, 1992, **30**, 1387–1390.

2.3 Summary of findings

The results of the FSA-based assays seemed to indicate that GlcNAc-SBn is a better substrate for MshB than GlcNAc or either of the other two chromogenic substrates (refer to Table 1 in publication, section 2.4). This was in agreement with the result obtained by Gammon *et al*. [8], which showed GlcNAc-SPh to be accepted as an alternative MshB substrate. However, when the chromogenic substrates were assayed by means of an end-point DTNB derivitization-based assay, MshB was found to have a high specific activity for GlcNAc-SDNP. No activity was observed with GlcNAc or the other two compounds (refer to Table 1 in publication, section 2.4). This led to the discovery of the S→N intramolecular rearrangement of the GlcNAc-SDNP deacetylated product, and ultimately to the development of the first continuous assay of MshB activity.

Unfortunately, GlcNAc-SDNP is not commercially available, and must be synthesized chemically. The yields for its preparation at present are still very low, and it also is poorly soluble in water which makes it difficult to handle. The next steps in the assay development would therefore be to identify chromogenic compounds that are more soluble in water, and to develop simpler, easier method to prepare these compounds.

2.4 References

1. Jothivasan VK, Hamilton CJ: Mycothiol: synthesis, biosynthesis and biological functions of the major low molecular weight thiol in actinomycetes. *Natural Product Reports.* 2008, **25**(6):1091-1117.
2. Anderberg SJ, Newton GL, Fahey RC: Mycothiol biosynthesis and metabolism. *Journal of Biological Chemistry.* 1998, **273**(46):30391-30397.
3. Huang X, Hernick M: A fluorescence-based assay for measuring N-acetyl-1-d-myo-inosityl-2-amino-2-deoxy-α-d-glucopyranoside deacetylase activity. *Analytical Biochemistry.* 2011, **414**(2):278-281.
4. Newton GL, Av-gay Y, Fahey RC: *N*-acetyl-1-d-myo-inosityl-2-amino-2-deoxy-α-D-glucopyranoside deacetylase (MshB) is a key enzyme in mycothiol biosynthesis. *Journal of Bacteriology.* 2000, **182**(24):6958-6963.
5. Newton GL, Ko M, Ta P, Av-Gay Y, Fahey RC: Purification and characterization of *Mycobacterium tuberculosis* 1D-myo-inosityl-2-acetamido-2-deoxy-α-D-glucopyranoside deacetylase, MshB, a mycothiol biosynthetic enzyme. *Protein Expression and Purification.* 2006, **47**(2):542-550.
6. Lamprecht DA: Development of a drug discovery protocol through the expression of key mycothiol biosynthetic enzymes from *Mycobacterium tuberculosis*. Stellenbosch: University of Stellenosch; 2008, pp. 62-94
7. Lamprecht DA, Muneri, N.O., Eastwood, H., Naidoo, K.J., Strauss, E. and Jardine M.A.: An enzyme-initiated Smiles rearrangement enables the development of an assay of MshB, the GlcNAc-Ins deacetylase of mycothiol biosynthesis. *Organic &Biomolecular Chemistry.* 2012, **10**:5278-5288.
8. Gammon DW, Steenkamp DJ, Mavumengwana V, Marakalala MJ, Mudzunga TT, Hunter R, Munyololo M: Conjugates of plumbagin and phenyl-2-amino-1-thioglucoside inhibit MshB, a deacetylase involved in the biosynthesis of mycothiol. *Bioorganic and Medicinal Chemistry.* 2010, **18**(7):2501-2514.

Chapter 3: Engineering of an α-N-acetylglucosaminidase for biocatalytic preparation of thioglycosides as alternative substrate of MshB

3.1 Introduction

The natural substrate of MshB is GlcNAc-Ins, a pseudo-disaccharide containing α(1→1) glycosidic linkage (Figure 2.1, and Section 2.1). GlcNAc-Ins is not commercially available, and it is therefore usually obtained by one of two ways: by multi-step chemical total synthesis [1-4], or by isolation from natural sources, such as from *M. smegmatis* [5]. However, these methods are not ideal because of the low yield of the final product that is obtained in both cases, and because of the time-consuming nature of obtaining GlcNAc-Ins in this manner. These drawbacks have led to numerous investigations and the characterization of various compounds containing the GlcNAc-backbone – such as GlcNAc, the GlcNAc *S*-bimane derivative MsmB and the GlcNAc thioglycosides used in the studies described in the previous chapter – as alternative substrates for MshB [6]. However, among these compounds the only commercially available alternative MshB substrate is the minimal substrate GlcNAc; all others must still be prepared by means of chemical synthesis. Unfortunately, MshB has been found to have both low activity and affinity for GlcNAc as compared to its natural or other alternative substrates, making it not ideal for use in the assaying of MshB activity [6-8].

In light of these shortcomings, the main focus of the studies described in this chapter was to engineer an α-*N*-acetylglucosaminidase enzyme to transform it into a biocatalyst that could be used to synthesize the natural or alternative substrates of MshB. These studies were performed in collaboration with Dr. Marco Moracci of the Institute of Protein Biochemistry (IPB) of the National Research Council (NRC) of Italy in Naples, and some of the results described here were obtained during a month-long research visit to his laboratory.

3.2 Strategy for the biocatalytic preparation of MshB substrate analogues

Various strategies have been developed for the preparation of biocatalysts that can be used for the synthesis of di- or oligosaccharides from suitably activated precursors. In all cases, these biocatalysts are based on naturally-occurring glycosidase (glycoside hydrolase) enzymes, which are subsequently mutated to introduce the desired reactivity. This section provides an introduction to these strategies.

3.2.1 Glycosidases and their catalytic mechanism

Glycosidases, also known as glycoside hydrolases, are catabolic enzymes that hydrolyze the glycosidic bond(s) of oligosaccharides or polysaccharides into their constituent sugar molecules [9]. They can act on both α- and β-glycosidic bonds, and can do so by endo-acting hydrolysis or exo-acting hydrolysis [10]. The resulting cleavage of the glycosidic bond happens with either net retention or net inversion of the anomeric configuration. It is this stereochemical course of the reaction that is used to classify the glycosidases as either inverting or retaining enzymes (Figure 3.1) [9-13].

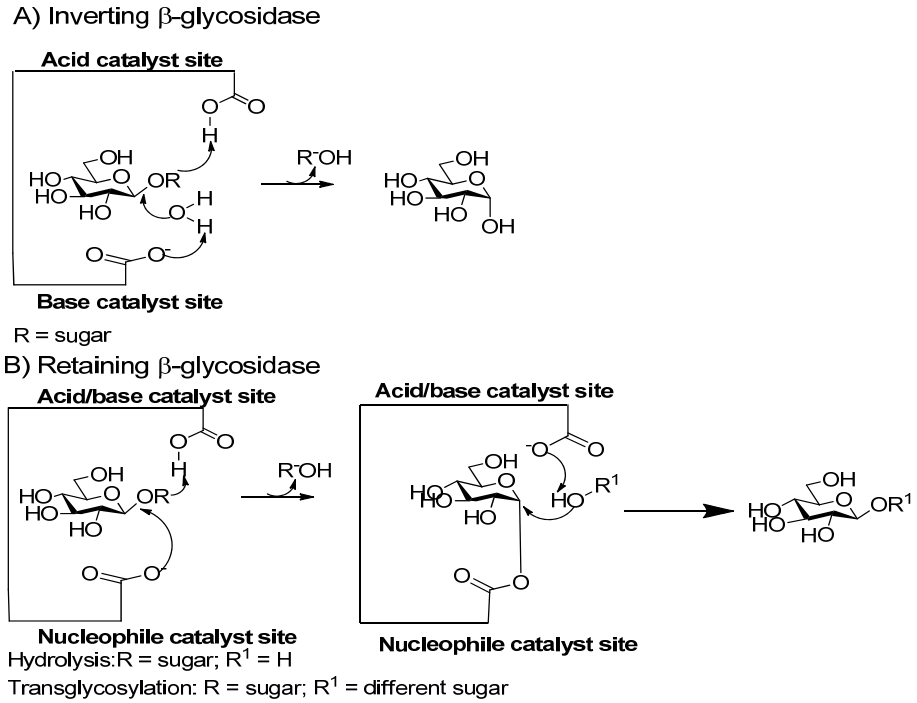


Fig. 3.1: Catalytic mechanisms of glycosidase. Figure adapted from Perugino *et al*. [13].

Inverting and retaining glycosidases share some common characteristics in their mechanisms and active site residues: they both contain an oxocarbenium ion-like transitions state and a pair of carboxylic acid-containing active site residues [9, 11]. However, inverting glycosidase enzymes use a direct displacement mechanism in which one active site residue acts as a base deprotonating the water molecule acting as a nucleophile, while the other residue acts as an acid that protonates the leaving group Figure 3.1A). Retaining glycosidases on the other hand utilizes a double displacement mechanism that consists of a glycosylation step in which the one sugar moiety is transferred to the enzyme, followed by deglycosylation which involves the hydrolysis of the glycosyl-enzyme intermediate through an oxocarbenium ion-like transition state to liberate the sugar molecules (Figure 3.1B) [9]. In this mechanism, the one carboxylic acid residue acts as a nucleophile that reacts with the anomeric carbon atom to form the covalent glycosyl-enzyme intermediate, while the other residue acts as an acid catalyst to cleave the glycosidic bond. This same residue subsequently acts as a base in the following attack by an incoming water molecule on the anomeric site to liberate the sugar molecule from the glycosyl-enzyme intermediate. This difference in mechanism is made possible by the active site residues of inverting glycosidases being far apart from each other compared to those of retaining glycosidases, which makes the intervention of a water molecule possible [11].

Glycosidases are also involved in an anabolic process called glycoside synthesis. Glycoside synthesis can occur either by one of two different mechanisms: by a reverse hydrolysis reaction, or by transglycosylation [13, 14]. Transglycosylation involves the replacement of water molecule with a different sugar molecule to promote synthesis [12]. However, the anabolic process of glycoside synthesis has its own limitations, as the end products of the reaction can act as substrates for the hydrolase activity of the enzyme, which results in low yield of the final product.

3.2.2 Converting glycosidase enzymes into glycosynthases

One strategy to overcome the problems associated with the anabolic process of glycoside synthesis is to mutate the glycosidase enzymes to transform them into glycosynthases. This is the first strategy that can be used to convert glycosidases into useful biocatalysts. Glycosynthases are prepared by mutating the active site

nucleophile residue of the glycosidase enzyme so that the first catalytic step of the hydrolysis reaction cannot take place [13, 15]. Such gycosynthases require a glycosyl donor with a good leaving group (such as fluoride, 2- or 4-nitrophenolate or 2,4-dinitrophenolate) to carry out the transglycosylation step (Figure 3.2).

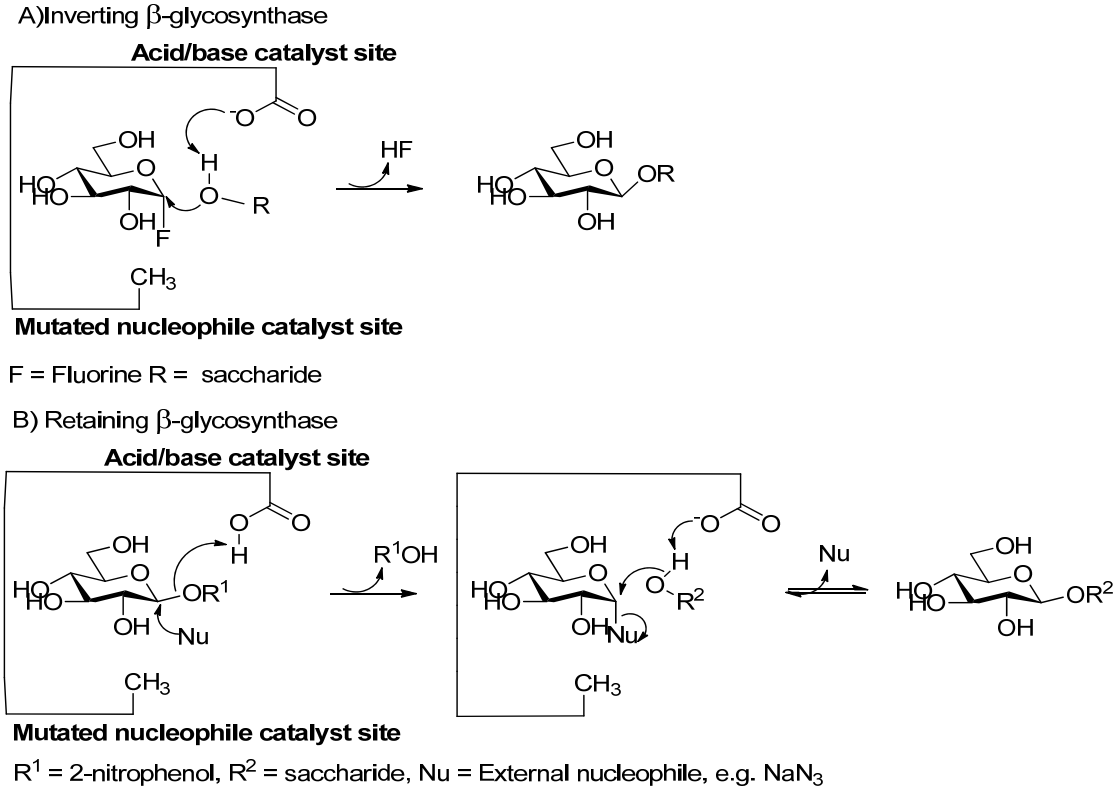


Fig. 3.2: The mechanism of β-glycosynthases generated from a β-retaining glycosidase. Figure adapted from Perugino *et al*. [13]

When an activated glycosyl donor is used with an anomeric configuration that is inverted relative to that of the normal substrate of the parent glycosidase, the glycosynthase acts as an inverting glycosynthase: the glycosyl donor acts as the intermediate of the reaction (Figure 3.2A) [13]. In order for the glycosynthase to act as a retaining glycosynthase, an activated glycosyl donor with the same anomeric configuration as the normal substrate must be used, as well as an additional external nucleophile, such as azide or formate (Figure 3.2B). In this case, the added external nucleophile mimics the role of the mutated nucleophile active site residue. This is followed by the attack of an acceptor molecule on the external nucleophile-substrate intermediate, which is assisted by the acid/base residue that now acts as a base.

The first glycosynthase was derived from the β-retaining glycosidase of *Agrobacterium faecalis* (*Agrobacterium* sp.) by Withers and co-workers in 1998 [16], wherein they mutated the nucleophile Glu^{358} to Ala. In an attempt to increase the activity of the glycosynthase, these researchers also mutated the nucleopile of an *Agrobacterium* sp. glycosidase (Glu^{358}) to Cys, Ser, Asp or Gly [15]. They found that Ser and Gly showed improved activity compared to the other mutations, with the Gly mutant showing the best activity overall. Since the discovery by Withers and co-workers [16] many glycosidases have been engineered using the same approach to obtain glycosynthases which are based on the retaining glycosidases. Most of the glycosynthases available to date are derived from β-retaining glycosidases of different microorganisms with only a few of α-retaining glycosidases modified to α-glycosynthase [13, 15, 17, 18]. The first α-glycosynthase was derived from the α-glucosidases of *Schizosaccharomyces pombe* by Chiba and co-workers in 2002 [19].

3.2.3 Converting glycosidase enzymes into thioglycoligases

A second strategy for the conversion of glycosidase enzymes into biocatalysts is based on the preparation of thioglycoligases. Thioglycoligases are formed by mutating the acid/base active site residue of a glycosidase to a non-catalytic residue, and using thiol-containing molecules as glycoside acceptors for the synthesis of thioglycosides [20-22]. Thioglycosides are glycans that contain a sulphur atom at the anomeric position instead of oxygen [21]. They are very stable in the presence of the glycosidase enzymes, as they are not natural substrates for the hydrolysis reaction catalysed by these enzymes [22]. While glycosynthases require either an activated substrate donor with an opposite anomeric configuration, or reactivation with an external nucleophile substrate, thioglycoligases only require a substrate donor with a good leaving group such as 2,4-dinitrophenolate (DNP) and a thiosugar molecule as an acceptor to carry out the thioglycoligase reaction (Figure 3.3) [20-23].

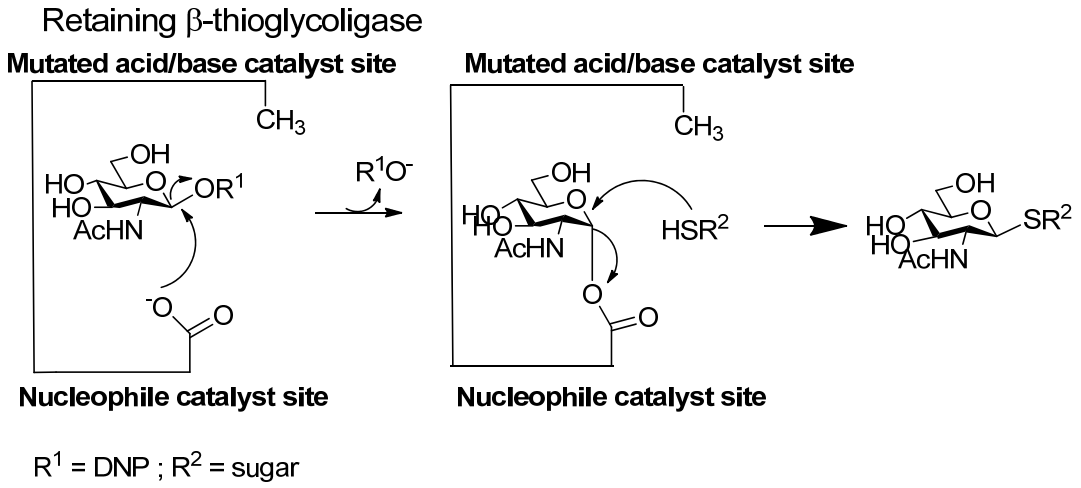


Fig. 3.3: The mechanism of thioglycoligase. Figure adapted from Perugino *et al*. [13].

Glycosidases were first modified to thioglycoligases by Withers and co-workers in 2003 [21]. They mutated the acid/base residue of the β-glycosidases from *Agrobacterium* sp. and *Cellulomonas fimi* (*C. fimi*)to Ala, and used 2,4-dinitrophenyl β-D-glucopyranoside and 2,5-dinitrophenyl β-D-mannopyranoside respectively for each enzyme as substrate donors, and *para*-nitrophenyl 4-deoxy-4-thio-β-D-mannopyranoside or *para*-nitrophenyl 4-deoxy-4-thio-β-D-xylopyranoside as acceptor molecules to demonstrate the concept of thioglycoligase-mediated synthesis. They incubated the mutants with the respective substrate donor and the acceptor molecule for a certain time at room temperature, and followed the reaction by thin layer chromatography (TLC) analysis of the reaction products, ^{1}H and ^{13}C NMR spectroscopy and mass spectrometry. The analyses showed the formation of disaccharide molecules that contained a sulphur-linked glycosidic bonds. However, when they used the parent sugar molecules of the wild type enzyme, i.e. *para*-nitrophenyl β-D-xylopyranoside and *para*-nitrophenyl β-D-glucopyranoside for *Agrobacterium* sp. and *C. fimi* respectively as acceptor molecules, no enzymatic reaction products was formed for the *Agrobacterium* sp. mutants. On the other hand *C. fimi* mutant was found to form a disaccharide molecule very slowly. This result confirmed that the glycosidases had been converted to thioglycoligases. To date four β-glycosidases and three α-glycosidases have been modified to thioglycoligases [17].

3.2.4 Converting a glycosidase into a suitable biocatalyst for the preparation of MshB substrate analogues

The α-*N*-acetylglucosaminidase from the Gram-positive anaerobic and heat-resistant endospore-forming bacterium, *Clostridium perfringens* (ATCC 13124), was identified as the potential candidate for the generation of a suitable biocatalyst for the preparation of MshB substrate analogue. The amino acid sequence of this enzyme, also known as *Cp*GH89, was found to be about 30% related to the human α-*N*-acetylglucosaminidase (NAGLU). Both *Cp*GH89 and NAGLU are classified as belonging to glycoside hydrolase family 89 (GH89), and act as retaining glycosidases that catalyze the hydrolysis of α-*N*-acetylglucosamine-containing oligosaccharides. The structure and mechanism of this enzyme was reported by Ficko-Blean *et. al*. [24] in 2008, wherein they showed that *Cp*GH89 consist of four distinct domains: an N-terminal domain, a linking α/β domain, a core $(\alpha/\beta)_8$ barrel and a C-terminal domain. The core $(\alpha/\beta)_8$ barrel (residues 280 to 620) was found to contain the active site which is in the shape of a sock formed by aromatic amino acids, with Glu483 acting as the acid/base residue and Glu^{601} as the nucleophilic active site residue.

Based on the previously described strategies for the conversion glycosidases into biocatalysts, two options were considered for the engineering of a biocatalyst that will enable the biocatalytic preparation of natural and/or alternative MshB substrates:

Option 1: Convert *Cp*GH89 into an α-glycosynthase for preparation of α-GlcNAc-containing glycosides.

In this approach an α-glycosynthase is generated from *Cp*GH89 by mutating the nucleophilic active site residue Glu^{601} to the non-reactive residues Gly, Ala or Ser. The generated α-glycosynthase can then be used to produce sugar molecules such as GlcNAc-Ins and the GlcNAc glycosides of 1,2,3-trihydroxycyclohexane and benzyl alcohol among others, by using either an inverting or retaining mechanism approach, as shown in figure 3.4. In the inverting mechanism, an activated β-glycosyl donor molecule will be used with the various acceptor molecules (R^1) to generate an α-glycosidic bond, while for the retaining mechanism an activated α-glycosyl donor molecule, external nucleophile (which mimics the mutated nucleophile) and various acceptor molecules will be utilized to form products with an α-glycosidic bond.

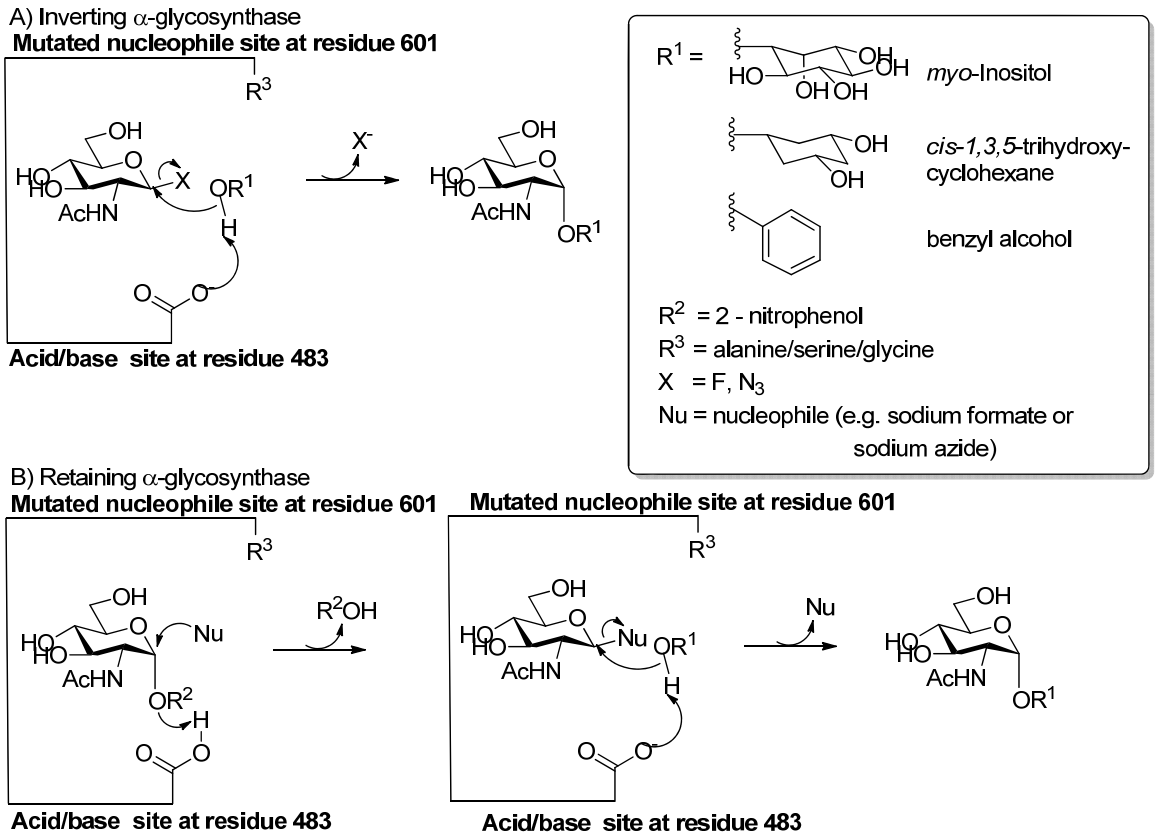


Fig. 3.4: The proposed approach for the synthesis of α-GlcNAc-containing glycosides using an α-glycosynthase. Panel A: For an inverting glycosynthase an activated β-glycosyl donor molecule is used together with various sugar acceptor molecules to retain the configuration of the α-glycosidic bond. Panel B: For a retaining glycosynthase, an activated α-glycosyl donor molecule is used together with various sugar acceptor molecules and an external nucleophile (to mimic the mutated nucleophile) to form products with an α-glycosidic bond.

The advantage of the using the glycosynthase approach is that it will allow for the preparation of the natural substrate of MshB (GlcNAc-Ins). However, it requires the use of either an activated β-GlcNAc glycosyl donor, which is difficult to chemically synthesize, for the inverting glycosynthase (Figure 3.1A), or the challenging reactivation of the glycosynthase with an external nucleophile if the α-retaining glycosynthase approach were to be used (Figure 3.1B).

Option 2: Convert *Cp*GH89 into an α-thioglycoligase as a biocatalyst for preparation of GlcNAc-containing thioglycosides.

In this approach α-thioglycoligases are generated from *Cp*GH89 by mutating the acid/base active site residue Glu^{483} to the non-reactive residues Gln, Ala or Ser. The generated α-thioglycoligases will be used to produce thiosugar molecules such as GlcNAc-SDNP and the GlcNAc thioglycosides of cyclohexanethiol, benzenethiol and *p*-nitrophenyl 4-thioglucopyranoside among others, by using a retaining mechanism approach as shown in figure 3.5. An activated α-glycosyl donor molecule together with various thio-acceptor molecules (R^2) will be used to form GlcNAc-based α-thioglycosides.

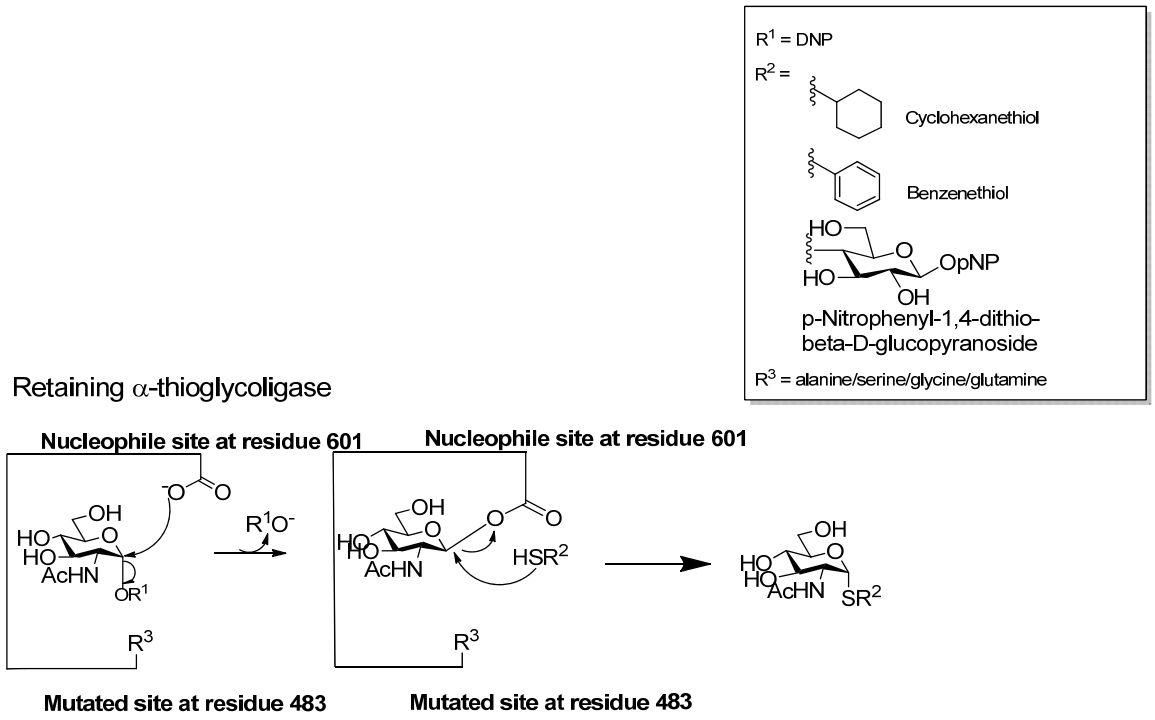


Fig. 3.5 The proposed approach to for the synthesis of α-GlcNAc-containing thioglycosides using an α-thioglycoligase. For a retaining thioglycosides, an activated α-glycosyl donor molecule is used together with various thiol-containing acceptor molecules to generate GlcNAc-based α-thioglycosides.

The benefit of the thioglycoligase approach is that an α-GlcNAc glycosyl donor is used; several such donors are commercially available. Furthermore, no external nucleophile is required, since the thiol group of the thiosugar acceptor reacts as the

nucleophile. However, a shortcoming is that thiosugars are not commercially available, and must be prepared through chemical synthesis.

For this study, it was decided to follow Option 2 and to prepare a thioglycoligase catalyst. However, attempts at preparing a glycosynthase (Option 1) were simultaneously initiated by our collaborator, Dr. Moracci of the IPB, Naples, Italy.

3.3 Results

3.3.1 Generation of plasmids encoding *Cp*GH89 mutants

An acid/base catalyst mutant library was generated by site-directed mutagenesis and single overlap extension (SOE) Polymerase Chain Reaction (PCR) using a plasmid containing the native *cpgh89* gene, which encodes the α-*N*-acetylglucosaminidase of *Clostridium perfringens*, as template. The acid/base catalyst site Glu483 was mutated to Gln, Ala and Ser residues. The expression vectors generated were named pET28a(+)-*Cp*GH89E483Q, pET28a(+)-*Cp*GH89E483A and pET28a(+)-*Cp*GH89E483S for the Glu, Ala and Ser mutants respectively. The modification was confirmed by DNA sequencing as shown in figure 3.6, wherein the conflict, which indicates the mutations, can clearly be seen: the genetic codes (GAA) which encodes for the Glu^{483} of the wild type was modified into CAA, GCA and TCA for generation of the various mutants.

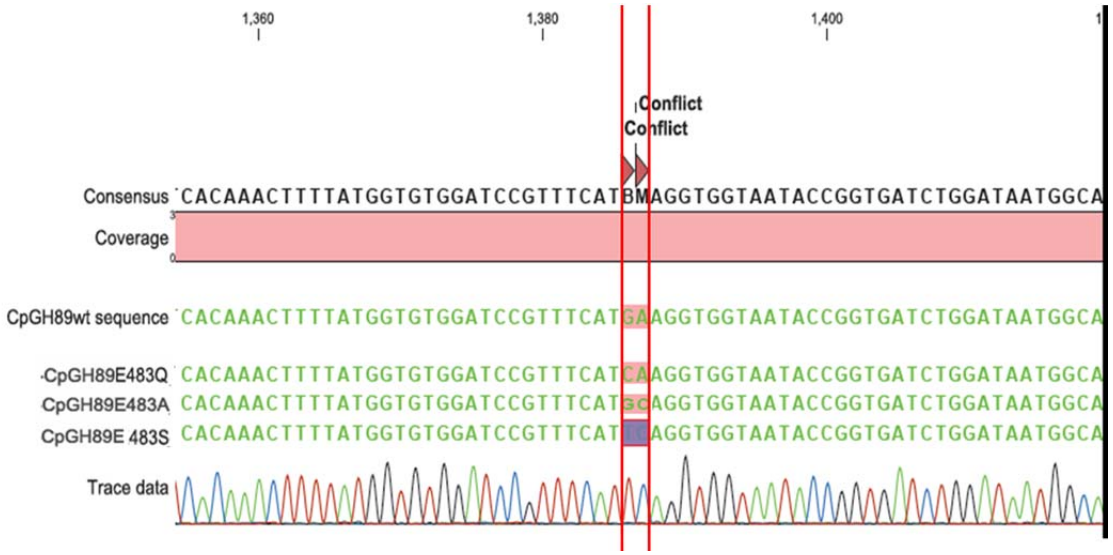


Fig. 3.6: DNA sequence alignments of the various mutants created from the wild type sequence. The genetic code GA is changed to CA, GC and TC for various mutants.

3.3.2 Expression and purification of wild type and mutant *Cp*GH89 proteins

The *Cp*GH89 enzymes were expressed from the pET28a(+) constructs, which encodes proteins with N-terminal 6xHis-tags, transformed into *Escherichia coli* (*E. coli*) BL21 (DE3). The expression conditions previously reported by Ficko-Blean *et al.* were used [24]. The proteins were purified by immobilized-metal affinity chromatography (IMAC) using an ÄKTA *prime* system. Cell crude extracts were applied to a 1 mL HiTrap Chelating HP column preloaded with Ni^{2+}, followed by stepwise elution using imidazole concentrations of between 25 and 150 mM (Figure 3.7A). The elution profile was monitored at 280 nm, which allowed the purification of the proteins to heterogeneity. The purity of the samples was confirmed by 12% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). A typical gel analysis of the heterologously expressed and purified proteins, which has a predicted molecular weight of 99 kDa, is shown in Figure 3.7B.

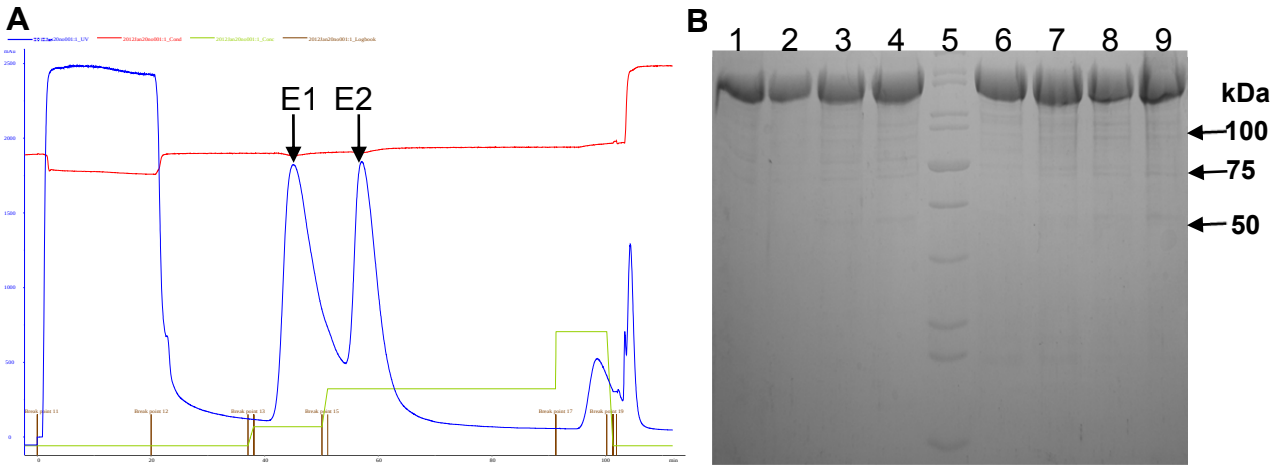


Fig. 3.7: Purification of the protein and nalyses of protein purification results by 12% SDS-PAGE.

A: Chromatography of the purification of proteins by IMAC using ÄKTA *prime* system following a stepwise elution using different concentration of imidazole. E1: Elution at 25 mM of imidazole and E2 elution at 150 mM of imidazole.

B: Analysis of purification by 12% SDS-PAGE. Lane 1-4 are proteins fraction eluted at E1: lane 1 *Cp*GH89, lane 2 *Cp*GH89E483Q, lane 3 *Cp*GH89E483A, lane 4 *Cp*GH89E483S and lane 6-9 are proteins fraction eluted at E2:lane 6 *Cp*GH89, lane 7 *Cp*GH89E483Q, lane 8 *Cp*GH89E483A, lane 9 *Cp*GH89E483S. Lane 5 is the PageRuler™ unstained protein marker.

3.3.3 Thermal stability of wild type and mutant *Cp*GH89 proteins

The secondary structure of *Cp*GH89 is made up of α-helix and β-sheets, and it contains four distinct domains (N-terminal, linking α/β, core $(\alpha/\beta)_8$ barrel and C-terminal) [24]. Replacement of the Glu^{483} with Gln, Ala and Ser might influence the secondary and tertiary structure of the protein, leading to misfolding. To verify if the modification of Glu^{483} did not influence the folding of the protein, thermal stability of the protein was tested using circular dichroism (CD), which monitors the secondary structure of the protein. The stability of the mutants was established by monitoring the temperature-dependent melting curves with CD at 220 nm. No differences in the stability of the protein were observed (Figure 3.8). All the mutants' protein had the same melting temperature of ± 70^0C, while the wild type started denaturing at about 3^0C higher temperature (± 73^0C) than the mutants. In addition, all the mutants and wild type protein seemed to have a structural conformation change around ± 53^0C.

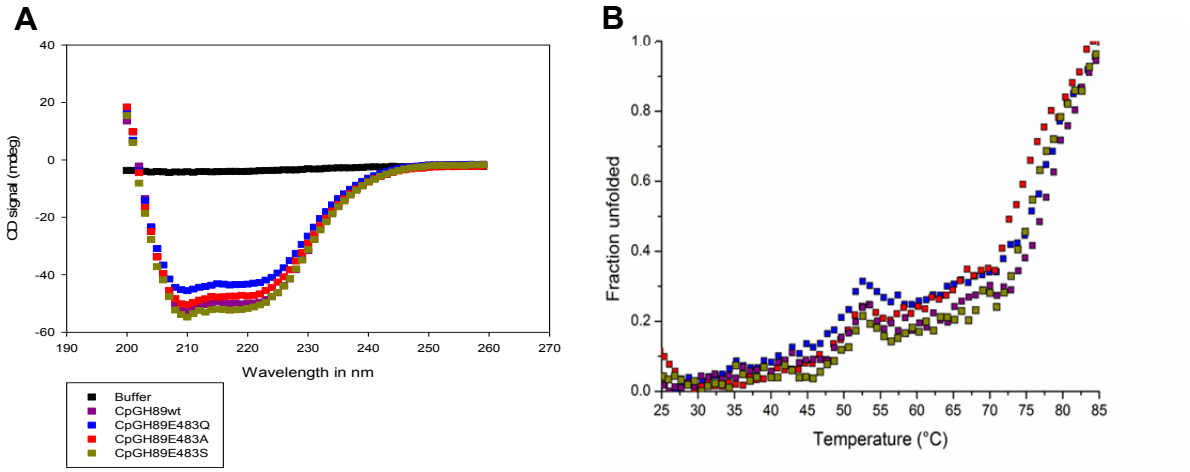


Fig. 3.8: The CD analysis of *Cp*GH89 and its mutants. **A**: CD spectra of the buffer, wild type and its mutants at 25^oC over a range of wavelengths (200-260nm). **B**: Melting curves of *Cp*GH89 and its mutants measured at 220 nm between 25-85^oC at an interval of 1 min per 1^oC change in temperature.

3.3.4 Hydrolytic activity of wild type and mutant *Cp*GH89 proteins

The α-glycosidase *Cp*GH89 utilizes the retaining mechanism for hydrolysis, with Glu601 acting as the nucleophile and Glu483 as the acid/base residue [24]. This means modification of any of the catalytic residues should result in the loss of hydrolysis activity. The hydrolytic activity of the three different *Cp*GH89 mutants and

the wild type enzyme were tested with two α-substrates: *p*-nitrophenyl α-*N*-acetyl-D-glucosaminide (*p*NP-α-GlcNAc) and *o*-nitrophenyl α-*N*-acetyl-D-glucosaminide (*o*NP-α-GlcNAc) (Figure 3.9). The hydrolysis of these compounds (*p*NP-α-GlcNAc and *o*NP-α-GlcNAc) by the enzyme liberate GlcNAc and 4-nitrophenol (*p*NP) or 2-nitrophenol (*o*NP). The liberation of *p*NP and *o*NP produces a yellow colour at high pH which can be detected spectrophotometrically at 400 nm, with extinction coefficients of 13300 $M^{-1}cm^{-1}$ and 1800 $M^{-1}cm^{-1}$ respectively in 0.1M sodium phosphate buffer at pH 7.3.

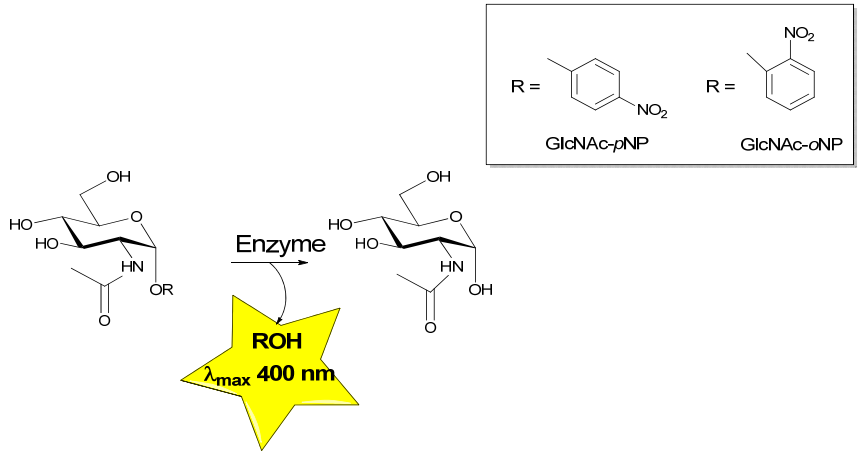


Fig. 3.9: The hydrolytic activity assay used to test the catalytic activity for *Cp*GH89 and its mutants.

The mutants' (*Cp*GH89E483Q, *Cp*GH89E483A and *Cp*GH89E483S) hydrolytic activity was found to be very low in comparison to the activity of the wild type enzyme as shown by the Michaelis-Menten plots in figure 3.10. The k_{cat}-values of the mutants were more than 10-fold smaller than that measured for the wild type enzyme, which clearly shows that these proteins have lost their hydrolytic activity (Table 3.1). This confirms the importance of the Glu483 for hydrolytic activity. The *Cp*GH89E483Q protein had the lowest turnover for substrates with k_{cat} = (6.2 ± 0.1)×10^{-3} sec^{-1} for GlcNAc-*p*NP and (3.3 ± 0.1)×10^{-2} sec^{-1} for GlcNAc-*o*NP. *Cp*GH89E483A and *Cp*GH89E483S showed similar turnover for the different substrates, with their k_{cat} values being higher than that shown by *Cp*GH89E483Q (Table 3.1).

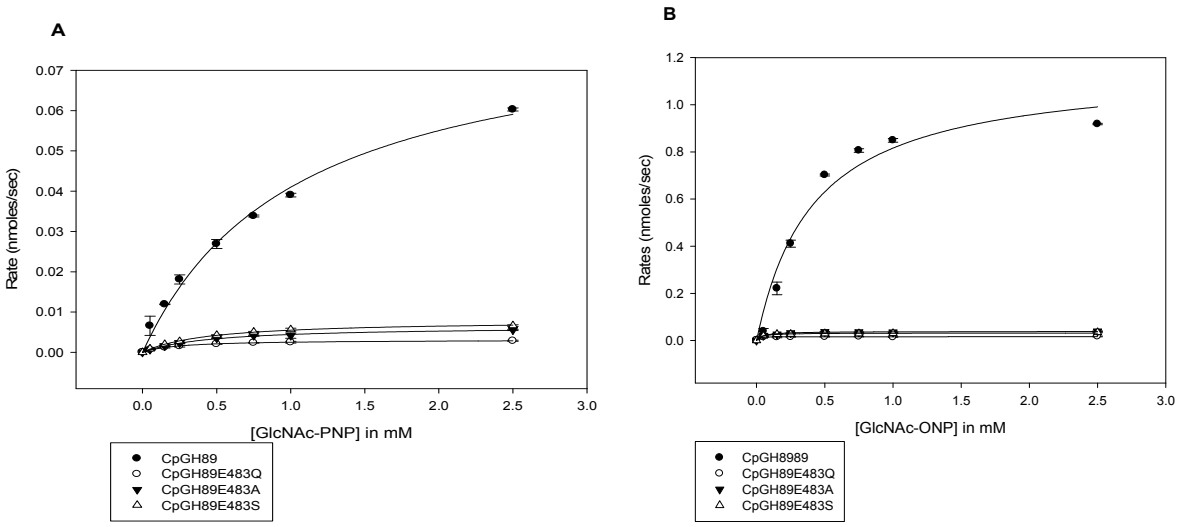


Fig. 3.10: Kinetic characterisation of the CpGH89 and its mutants. The graphs were plotted using non-linear fitting of the Michaelis-Menten equation (V_{max}=*kcat*.[S]/K_M+[S]) to the obtained data with the best-fit represented by the solid lines. Each point was done in triplicate with error bars. Panel **A**: Kinetic characterization of *Cp*GH89 and its mutants with GlcNAc-*p*NP as substrate. Panel **B**: Kinetic characterization of CpGH89 and its mutants with GlcNAc-*o*NP as substrate.

Table 3.1: The kinetic parameters of the *Cp*GH89 wild type and its mutants determined with GlcNAc-*p*NP and GlcNAc-*o*NP as substrate

Substrates & enzymes used	**K_m (mM)**	**V_{max} (nmoles/sec)**	**k_{cat} (sec^{-1})**	**k_{cat}/K_m ($mM^{-1} sec^{-1}$)**
GlcNAc-*p*NP				
*Cp*GH89wt	1.0 ± 0.1	$(8.4 \pm 0.3)\times10^{-2}$	$(1.7 \pm 0.1)\times10^{-1}$	$(1.6 \pm 0.8)\times10^{-1}$
*Cp*GH89E483Q	$(2.4 \pm 0.2)\times10^{-1}$	$(3.1 \pm 0.1)\times10^{-3}$	$(6.2 \pm 0.1)\times10^{-3}$	$(2.6 \pm 0.7)\times10^{-2}$
*Cp*GH89E483A	$(4.6 \pm 0.4)\times10^{-1}$	$(6.5 \pm 0.2)\times10^{-3}$	$(1.3 \pm 0.0)\times10^{-2}$	$(2.8 \pm 0.9)\times10^{-2}$
*Cp*GH89E483S	$(4.4 \pm 0.3)\times10^{-1}$	$(7.9 \pm 0.2)\times10^{-3}$	$(1.6 \pm 0.0)\times10^{-2}$	$(3.6 \pm 1.6)\times10^{-2}$
GlcNAc-*o*NP				
*Cp*GH89wt	$(4.2 \pm 0.6)\times10^{-1}$	1.2 ± 0.1	2.3 ± 0.1	5.6 ± 2.0
*Cp*GH89E483Q	$(1.1 \pm 0.5)\times10^{-2}$	$(1.6 \pm 0.0)\times10^{-2}$	$(3.3 \pm 0.1)\times10^{-2}$	3.1 ± 0.2
*Cp*GH89E483A	$(4.6 \pm 0.5)\times10^{-2}$	$(3.8 \pm 0.0)\times10^{-2}$	$(7.6 \pm 0.1)\times10^{-2}$	1.7 ± 0.0
*Cp*GH89E483S	$(3.4 \pm 0.5)\times10^{-2}$	$(3.2 \pm 0.0)\times10^{-2}$	$(6.4 \pm 0.1)\times10^{-2}$	1.9 ± 0.3

3.3.5 Evaluation of the thioglycoligase ability of the mutants

To test whether the various mutants can promote transglycosylation, 2,4-dinitrophenyl α-*N*-acetyl-D-glucosaminide (DNP-α-GlcNAc) was used as the donor substrate, while sodium azide (NaN_3) was used as the nucleophile to mimic the role of the thiosugar. NaN_3 was used for this first test to determine if the active site can accommodate an external nucleophole, and because no thiosugar is commercially available, and will have to be synthesized first before it can be used in such tests. DNP-α-GlcNAc was chosen as the substrate donor because the pK_a value of its leaving group (2,4-dinitrophenolate, DNP) is low (3.96) compared to those of *p*NP-α-GlcNAc and *o*NP-α-GlcNAc (~pK_a 7.18), indicating that it should have a better leaving group ability. DNP-α-GlcNAc was not used to test for the hydrolytic activity of the mutants because it is not commercially available and must be prepared by chemical synthesis. For this study, it was prepared by Dr. Emiliano Bedini, collaborator of Dr. Moracci.

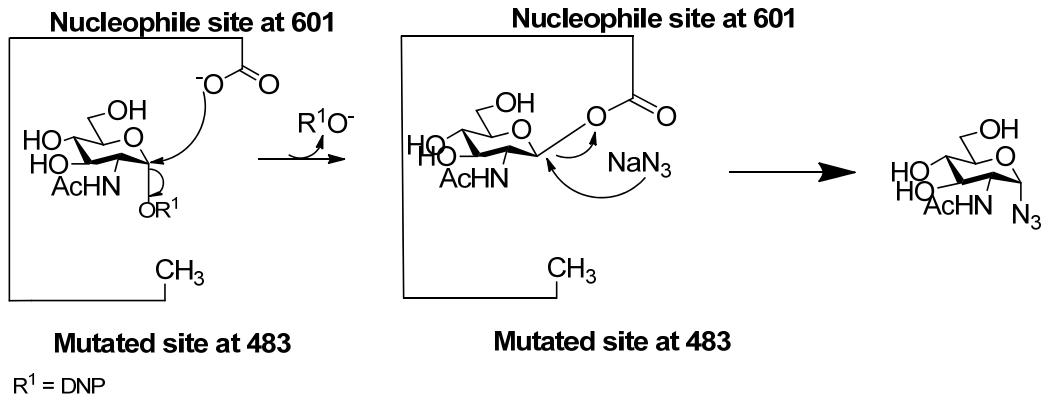


Fig. 3.11: The catalytic mechanism of the retaining α-thioglycoligases, *Cp*GH89 mutant enzymes. Mutation of the acid-based residue allows the glycosyl-enzyme intermediate to be formed by reaction of the active site nucleophile with the activated donor DNP-α-GlcNAc. If the intermediate is formed, and external nucleophile such as NaN_3 can react with it to demonstrate its potential to react with thiosugars.

DNP-α-GlcNAc (5 mM), 0.017-0.275 mg/mL of mutations and 100 mM of NaN_3 in 100 mM sodium phosphate buffer, pH 7.3, were incubated at 37^0C overnight followed by TLC analysis of the reaction products. The TLC analysis revealed the formation of at least three distinct products for the reactions where the mutant enzymes, *Cp*GH89E483Q, *Cp*GH89E483A and *Cp*GH89E483S that is R4+, R5+ and R6+ respectively in Figure 3.11, were incubated in the presence of NaN_3. Only two

products were observed following the reaction for the reactions with the wild type enzyme and the reactions with the other mutants in the absence of NaN_3 (R1-, R2-, R3-, R4-, R5-,R6-) and present of NaN_3 (R1+, R2+ and R3+) , and only one compound (single spot) was observed for the negative controls reaction (R7+ and R7- in Figure 3.11). The products spot for R4+, R5+ and R6+ with the highest retardation factor (R_f) of 0.88 was only visible under ultraviolet (UV) light, while the other two products spot, Rf 0.77 and 0.32, were visible when stained with 10% sulphuric acid in methanol. The products spot with Rf 0.88 is 2,4-dinitrophenol, while the middle R_f-value 0.77 had more or less the same Rf-value measured for STD2 (β-GlcNAc-N_3) (Figure 3.11). The R_f-value of the bottom spot was below to that of GlcNAc (STD1).

The products spot with Rf 0.77 was only observed in the reactions with NaN_3 which were catalysed by *Cp*GH89E483Q, *Cp*GH89E483A and *Cp*GH89E483S mutants. This indicates that these proteins are able to promote transglycosylation in the presence of NaN_3. No transglycosylation reaction was observed in the absence or presence of NaN_3 for reactions catalysed by *Cp*GH89E483G and *Cp*GH89E601G (*Cp*GH89E601G mutant is a potential α-glycosynthase created in Dr. M. Moracci lab by mutating the nucleophilic active site residue Glu^{601} to Gly, and was added to this test for comparison). The observed products spots for both these enzymes had the same R_f-values as the ones obtained for the reactions catalysed by the wild type enzyme. This could mean that the replacement of Glu with Gly promotes folding of the protein in such a way that another amino acid residue is revealed which then acts as a nucleophile promoting hydrolysis. However, for the other α-thioglycoligases (in which the acid/base residue was mutated to residues other than Gly) this does not occur, indicating that these enzymes have the potential to be used as biocatalysts.

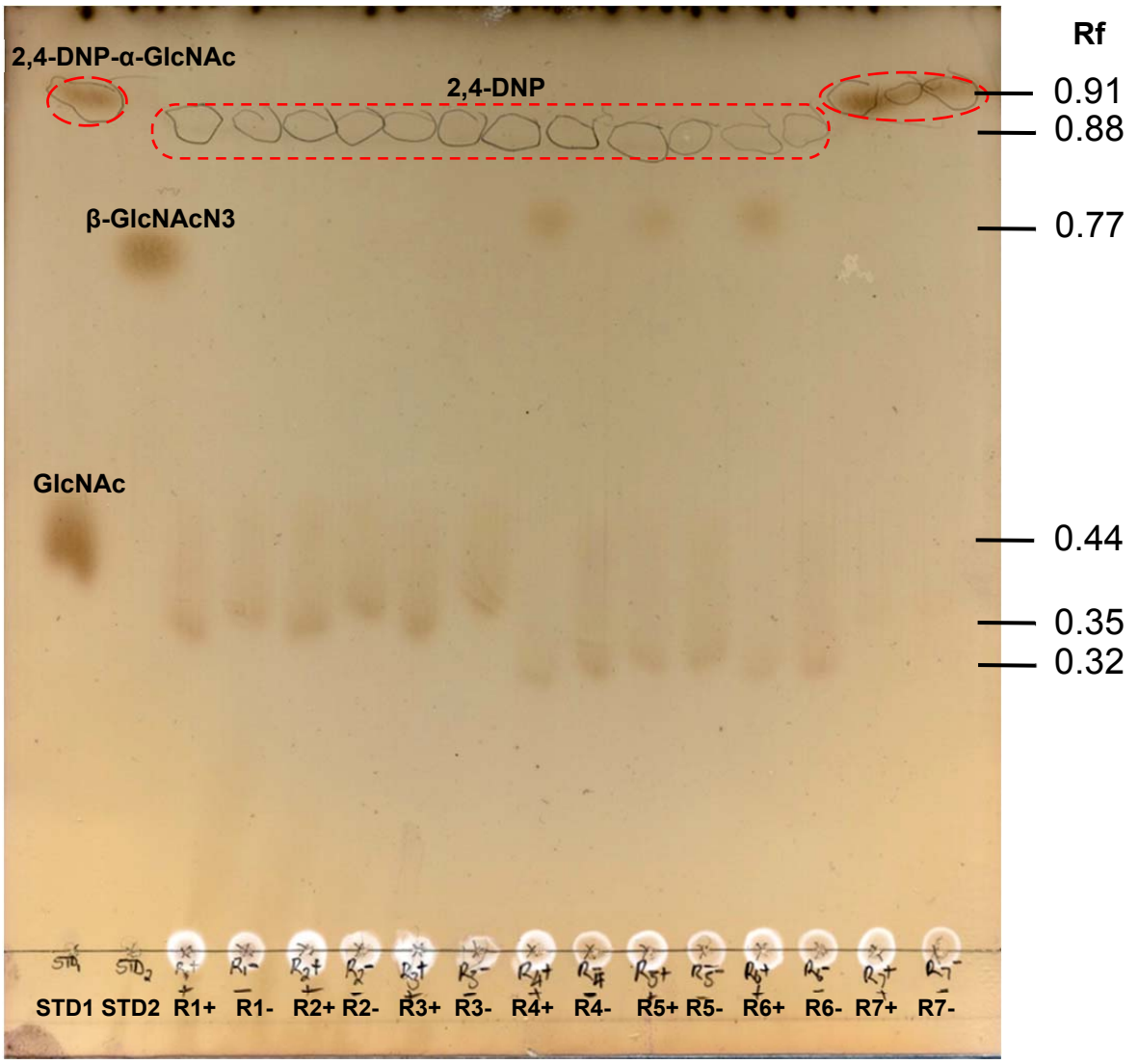


Fig. 3.11: TLC analysis of the thioglycoligase ability of the *Cp*GH89 mutants. The lanes are identified as follows: STD1, 10 µL of 20 mM GlcNAc and 5µL of 30 mM DNP-α-GlcNAc; STD2, 5 µL of 20 mM β-GlcNAcN_3;

R1+/-, *Cp*GH89 wild type with and without NaN_3;

R2+/-, *Cp*GH89E601G with and without NaN_3;

R3+/-, *Cp*GH89E483G with and without NaN_3;

R4+/-, *Cp*GH89E483Q with NaN_3 and without NaN_3;

R5+/-, *Cp*GH89E483A with and without NaN_3;

R6+/-, *Cp*GH89E483S with and without NaN_3;

R7+/-, no enzyme with and without NaN_3 (negative control).

The red circles indicate the products/reactants spots visualised by UV light, while the brown products/reactants spots were visible after staining with 10% sulphuric acid in methanol and subsequent charring.

3.4 Discussion

The use of thioglycosides as therapeutic drugs instead of glycosides has increasingly been investigated over the past few years [21]. However, to produce thioglycosides a new enzymatic approach for their synthesis is needed. This has led to the engineering of several glycosidases to thioglycoligases through the removal of the acid-base catalytic residue by site-directed mutagenesis [17, 21], with most of the thioglycoligases created to date being modified from β-retaining glycosidases from microorganisms such as *Agrobacterium sp., C. fimi and Bacillus circulans* [17].

In this chapter the cloning, heterologous expression and IMAC purification of new potential α-thioglycoligases from a retaining α-glycosidase, the α-*N*-acetlyglucosaminidase of *C. perfringens*, were described. Three potential α-thioglycoligases were generated through site-directed mutagenesis of the acid/base residue (Glu^{483}) using SOE PCR and subsequent expression from the resulting plasmids. Previously identified expression parameters that were used to produce the wild type allowed for over-expression of the mutant proteins in large quantity [24, 25]. The resulting enzymes were found to have a thermal stability similar to that of the wild type enzyme, indicating that the mutations did not influence the overall structure or folding of the mutant proteins (Figure 3.8).

2,4-Dinitrophenol-containing donors have been reported to be better substrates for thioglycoligase reactions than the corresponding *p*-nitrophenol or *o*-nitrophenol substrates due to the better leaving group ability of 2,4-dinitrophenolate [23]. In the case of the *Cp*GH89 mutants, an ideal situation would have been to conduct all tests using DNP-α-GlcNAc. However, since it is not commercially available and it is produced in low yield by chemical synthesis, the residual hydrolytic activity of the mutants could not be tested with this compound as a substrate. Nevertheless, the hydrolytic activities of the mutants were measured using *p*NP-α-GlcNAc and *o*NP-α-GlcNAc. The *Cp*GH89 mutants' hydrolytic activity rates were found to be less than 95-fold as compared to that of the wild type enzyme (Figure 3.10). These results confirm those obtained by Ficko-Blean *et. al*, [24] in which they discovered the catalytic active site residue of *Cp*GH89 to be Glu^{601} and Glu^{483}.

The generated α-thioglycoligases (*Cp*GH89E483Q, *Cp*GH89E483A and *Cp*GH89E483S) were found to be good candidates to use as α-thioglycoligase for the

synthesis of α-thioglycosides. Transglycosylation was promoted by the α-thioglycoligases in the presence of NaN_3 (Figure 3.11). This shows that CpGH89E483Q, CpGH89E483A and CpGH89E483S are able to catalyse the formation of the glycosyl-enzymes intermediate (glycosylation step) which is essential for deglycosylation to take place. However, it was quite surprising that the CpGH89E483G mutant did not show transglycosylation, but instead only showed hydrolysis, both in presence and absence of NaN_3. This could indicate that NaN_3 is unable to resolve the glycosyl-enzyme intermediate, or that the mutation is interfering with the formation of glycosyl-enzyme intermediate.

3.5 Conclusion

In this study the first examples of α-thioglycoligases from an α-*N*-acetylglucosaminidase of *C. perfringens* were engineered by mutation of the enzyme's acid/base active site residue. All three enzymes were found to be good candidates for the enzymatic synthesis of α-thioglycosides which can be used as alternative substrate for MshB. Future studies include chemical synthesis of thiosugar molecules as acceptor in order to validation the α-thioglycoligase ability and the implementation of these enzymes for the large production of a range of MSH analogs. These analogs will be evaluated as potential alternative substrate for MshB. In addition they will also be evaluated as potential alternative substrate/inhibitors compounds against the enzymes involved in the MSH biosynthesis pathway and also against the MSH-dependent enzymes.

3.6 Experimental procedures

All the chemicals used and columns for the ÄKTA*prime* system were purchased from Sigma-Aldrich unless stated otherwise. All restriction enzymes and the Eppendorf Perfectprep® Gel Cleanup kit were purchased from Fermentas. The DNA primers, DNA sequencing and Zyppy™ plasmid Miniprep Kit (Zymo Research) was purchased from Inqaba Biotechnical Industries (Pty) Ltd. The expression plasmid were obtained from Novagen, while the ligation kit was from New England BioLabs. The pET28a(+)-*Cp*GH89 plasmid, containing the optimized DNA sequence of α-*N*-acetylglucosaminidase, was obtained from Dr. Marco Moracci.

3.6.1 Generation of the various thioglycoligases

3.6.1.1 Site-direct mutagenesis

The pET28a(+)-*Cpgh89* plasmid was used as DNA template in all PCR reactions. Single overlap extension (SOE) Polymerase Chain Reaction (PCR) was used to generate the site mutation into *Cpgh89* gene. The primers used to generate the three different isoforms are listed in Table 3.2.

The following PCR conditions were used to generation the DNA fragments consisting of the mutants; 10x *Pfu* buffer without $MgSO_4$, 1.25 U of *Pfu* DNA polymerase, between 2-4 mM of $MgSO_4$, 0.4 mM of dNTPs cocktail, 1 ng of the template DNA (pET28a(+)-*Cp*gh89 optimized plasmid), 1 mM of non-mutagenic/mutagenic forward primer and 1 mM of non-mutagenic/mutagenic reverse primer in a total volume of 25 µL. The void volume was filled with sterilized distilled water. The thermal cycling conditions used are shown in Table 3.3.

Table 3.2: DNA sequence of the primers used to generate various mutants of *Cp*GH89

Name		Primer
Mutagenic primers	*Cp*GH89E483Q	5' GGATCCGTTTCAT**CAA**GGTGGTAATACCGG-3'-fw
		5'-CCGGTATTACCAC**CTT**GATGAAACGGATCC-3'-rev
	*Cp*GH89E483A	5'-GGATCCGTTTCAT**GCA**GGTGGTAATACCGG-3'-fw
		5'-CCGGTATTACCAC**CTG**CATGAAACGGATCC-3'-rev
	*Cp*GH89E483S	5'-GGATCCGTTTCAT**TCA**GGTGGTAATACCGG-3'-fw
		5'-CCGGTATTACCAC**CTG**AATGAAACGGATCC-3'-rv
Non-mutagenic primers	T7 promoter	5'-TAATACGACTCACTATAGGG-3'
	T7 terminator	5'-GCTAGTTATTGCTCAGCGG-3'

Table 3.3: Thermal cycling conditions for PCR

Step	Temperature in ^{0}C	Time in seconds	Number of cycles
Initialization	94	120	1
Denaturation	94	30	30
Annealing	45	30	
Elongation	68	150	
Final elongation	68	360	1
Final hold	4	Infinite	None

The PCR products were analysed by gel electrophoresis (1% agarose gel) stained with SYBR® safe and compared to marker (2-log ladder marker) to identify the correct size band. The most concentrated bands of the correct size (1754bp) were purified from the gel using the Eppendorf Perfectprep® Gel Cleanup kit. The purified PCR products were amplified again using the non-mutagenic primers (T7 primers) to link the DNA fragments in the same manner as described above. The PCR products (2703bp) were analysed and purified as before.

3.6.1.2 Cloning of the PCR products into the plasmid pET28a(+) vector

Restriction digest of the PCR products and plasmid. Fastdigest *Nhe*I and *Xho*I restriction enzymes, were used to generate sticky ends in pET28a(+) and the PCR products using the following digestion reaction conditions: 1 U of *Nhe*I, 1 U of *Xho*I and 1× Fastdigest buffer and DNA fragments/plasmid to a total volume of 20 µl. The reaction mixtures were incubated at 37^{o}C for 1 hour. The digested products were visualized by gel electrophoresis and purified as before.

DNA ligation and screening. All constructs were cloned into pET28a(+) (Novagen) which contains an *N*-terminal 6×His-tag and a T7 promoter for heterologous protein expression in *E. coli*. The PCR products to plasmid recombinant ratio were 1:1. For the ligation reaction 10 U of T4 DNA Quickligase (New England BioLabs) in 1× Quick ligase buffer was added to the predetermined amount of PCR product and plasmid and incubated at 25^{o}C for 20 minutes. The ligation reactions were transformed as is into chemically competent *E. coli* MachI cells, plated on Luria Bertani agar media containing 30 mg/L kanamycin and incubated at 37^{o}C overnight. The insertion of the PCR products into the plasmids pET28a(+) vector was verified by screening the colonies using the Colony Fast Screen kit from Epicentre wherein a smear of a colony was added to 25 µL of Epilyse solution 1 (30 mM Tris-HCl, pH 8.0, 5 mM Na_2EDTA, 50 mM NaCl, 20 % sucrose, 0.047 µg.mL^{-1} lysozyme and 0.047 µg.mL^{-1} RNAse). After the addition 10 µL of Epilyse solution 2 (1× TAE, 2% SDS, 5% sucrose and 2 mg.mL^{-1} bromophenol blue) the mixture was analysed by gel electrophoresis as before in comparison to undigested parent pET28a(+) plasmid (plasmids in which the PCR products have been inserted should be larger than the parents plasmid). Positive screening results were confirmed by diagnostic digests of the constructs with

the restriction enzymes used to generate the sticky ends for ligation and by sequencing.

3.6.2 Expression of *Cp*GH89 wild type and mutants

All expressions were done from *E. coli* BL21 (DE3). The expressing conditions were based on the procedure used by Ficko-Blean *et al.* [25], but with minor modifications. The cultures were grown until the optical density at 600nm (OD_{600}) reached 0.600. After induction at 37°C with 0.33 mM isopropyl-1-thio-β-D-galactropyranoside (IPTG), the cultures were grown overnight before being harvested by centrifugation at 17600×*g* for 20 minutes at 10°C. The pellets cells were stored at -20°C until use.

3.6.3 Protein purification

The cells were resuspended in binding buffer (10 mL of buffer for every gram of dry pellet) consisting of 20 mM Tris-HCl, pH 7.9, 5 mM imidazole, 500 mM NaCl, and 0.05 % ($^{w}/_{v}$) NaN_3 and lysed by sonication. The pellet debris was pelleted by centrifugation at 75000×g for 20 minutes at 10°C and the lysate was filtered with a Acrodisc PSF Gx/GHP 0.45 µm filter (Pall) prior to injection into the ÄKTA *prime* system for IMAC purification. The supernatant was loaded onto a 1 mL HiTrap Chelating HP column preloaded with Ni^{2+}. The bound proteins were eluted by using a His-tagged elution buffer (20 mM Tris-HCl, pH 7.9, 500 mM imidazole, 500 mM NaCl, and 0.05% ($^{w}/_{v}$) NaN_3) in a step elution between 25 and 150 mM of imidazole concentrations. The elution profile was followed by UV at 280 nm. The high salt concentration from the eluted proteins was removed by buffer exchange on the on the ÄKTA *prime* system using a 5 mL HiTrap desalting column and gel filtration buffer consisting of 25 mM Tris-HCl, pH 8.0 and 5 mM $MgCl_2$. The proteins were stored at -80°C in 5 % glycerol until use. The purity of the proteins was confirmed by SDS-PAGE analysis.

3.6.4 Protein quantification

The concentrations of the proteins were quantified by using the Bradford protein assay using Bovine serum albumin (BSA) as standard. The amount of Coomassie Brilliant Blue dye bounded to the BSA standards and the unknown protein was measured at 595 nm using the Thermo Varioskan™ after incubation at room

temperature for 15 min. The processing of the data was done using Thermo Varioskan™ Skanit software.

3.6.5 Hydrolytic Activity assay of the mutants

All the activity assays with the mutated enzymes were conducted with two different substrates, *p*-nitrophenyl α-*N*-acetyl-D-glucosaminide (*p*NP-α-GlcNAc) and *o*-nitrophenyl α-*N*-acetyl-D-glucosaminide (*o*NP-α-GlcNAc), by means of a microplate high-throughput continues spectrophotometric assay. All reactions were conducted in triplicate at 25°C in a finale reaction volume of 250 µL. The assays contained 2 µM enzyme, 100 mM sodium phosphate buffer, pH 7.3, and with a range of substrate concentrations varying from 0-2.5 mM. The liberation of the *p*-nitrophenyl and *o*-nitrophenyl was monitored at 400 nm for 10 minutes. The rates were calculated using linear regression curve and the Michealis-Mention parameters determined. The extinction coefficient of ε = 13.3 $mM^{-1}.cm^{-1}$ for *p*-nitrophenyl and ε = 1.8 $mM^{-}1.cm^{-1}$ for *o*-nitrophenyl was used to calculate the concentrations using the Beer-Lambert law.

3.6.6 Evaluation of the thermal stability of α-*N*-acetlyglucosaminidase and its mutants

CD spectroscopy and thermal analysis was conducted using a Photophysics Chirascaw-plus spectrometer. The samples contained 0.5 mg/mL protein in Tris buffer (25 mM Tris-HCl, pH 8 and 5 mM $MgCl_2$. The optimum wavelength for analysis of thermal stability was determined by a CD scan from 200-300 nm at 25°C and the thermal stability experiments was conducted by heating the samples from 25°C to 100°C and following the denaturation of the protein at 220nm.

3.6.7 Evaluation of the thioglycoligase ability of the mutants

The evaluation reactions of the thioglycoligase ability were conducted at the laboratory of Dr Marco Moracci at IPB-CNR in Naples, Italy.

The three mutants described above were tested for their thioglycoligase ability together with other two α-N-acetylglucosaminidase mutants (*Cp*GH89E601G and *Cp*GH89E483G) provided by Dr. M. Moracci. DNP-α-GlcNAc was used as the donor substrate. All thioglycoligase reactions were done in 100 mM sodium phosphate buffer, pH 7.3, with 0.017-0.275 mg/mL enzyme and 5 mM of DNP-α-GlcNAc in a

final volume of 200 µl. The reactions were done in the presence and absence of 100 mM NaN_3. After incubation at 37^{o}C overnight, the reaction mixtures were analyses by thin layer chromatography (TLC) (Ethyl acetate: methanol: water, 7:2:1), visualised with UV light to observe the nitrophenol group and 10% sulphuric acid in methanol. For each reaction 4x5 µL of the solution was spotted on the TLC.

3.7 References

1. Ajayi K, Thakur VV, Lapo RC, Knapp S: Intramolecular α-glucosaminidation: synthesis of mycothiol. *Organic Letters.* 2010, **12**(11):2630-2633.
2. Jardine MA, Spies HSC, Nkambule CM, Gammon DW, Steenkamp DJ: Synthesis of mycothiol, 1D-1-O-(2-[N-acetyl-l-cysteinyl]amino-2-deoxy-α-d-glucopyranosyl)-myo-inositol, principal low molecular mass thiol in the actinomycetes. *Bioorganic and Medicinal Chemistry.* 2002, **10**(4):875-881.
3. Knapp S, Gonzalez S, Myers DS, Eckman LL, Bewley CA: Shortcut to mycothiol analogues. *Organic Letters.* 2002, **4**(24):4337-4339.
4. Nicholas GM, Eckman LL, Kovac P, Otero-Quintero S, Bewley CA: Synthesis of 1-d- and 1-l-myo-inosityl 2-N-acetamido-2-deoxy-alpha-d-glucopyranoside establishes substrate specificity of the *Mycobacterium tuberculosis* enzyme AcGI deacetylase. *Bioorganic and Medicinal Chemistry.* 2003, **11**(12):2641-2647.
5. Newton GL, Av-gay Y, Fahey RC: N-Acetyl-1-d-myo-Inosityl-2-Amino-2-Deoxy-α-D-glucopyranoside deacetylase (MshB) is a key enzyme in mycothiol biosynthesis. *Journal of Bacteriology.* 2000, **182**(24):6958-6963.
6. Newton GL, Ko M, Ta P, Av-Gay Y, Fahey RC: Purification and characterization of *Mycobacterium tuberculosis* 1d-myo-inosityl-2-acetamido-2-deoxy-α-d-glucopyranoside deacetylase, MshB, a mycothiol biosynthetic enzyme. *Protein Expression and Purification.* 2006, **47**(2):542-550.
7. Huang X, Hernick M: A fluorescence-based assay for measuring N-acetyl-1-d-myo-inosityl-2-amino-2-deoxy-α-d-glucopyranoside deacetylase activity. *Analytical Biochemistry.* 2011, **414**(2):278-281.
8. Lamprecht DA, Muneri, N.O., Eastwood, H., Naidoo, K.J., Strauss, E. and Jardine M.A.: An enzyme-initiated Smiles rearrangement enables the development of an assay of MshB, the GlcNAc-Ins deacetylase of mycothiol biosynthesis. *Organic and Biomolecular Chemistry.* 2012, **10**:5278-5288.
9. Withers SG: Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers.* 2001, **44**(4):325-337.
10. Lieshout JFTv: Characterization and engineering of thermostable glycoside hydrolases. Wageningen: Wageningen University; 2007.
11. Wang Q, Graham RW, Trimbur D, Warren RAJ, Withers SG: Changing enzymic reaction mechanisms by mutagenesis: Conversion of a retaining

glucosidase to an inverting enzyme. *Journal of the American Chemical Society.* 1994, **116**(25):11594-11595.

12. Kittl R, Withers SG: New approaches to enzymatic glycoside synthesis through directed evolution. *Carbohydrate Research.* 2010, **345**(10):1272-1279.
13. Perugino G, Trincone A, Rossi M, Moracci M: Oligosaccharide synthesis by glycosynthases. *Trends in Biotechnology.* 2004, **22**(1):31-37.
14. Crout DHG, Vic G: Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. *Current Opinion in Chemical Biology.* 1998, **2**(1):98-111.
15. Hancock SM, Vaughan MD, Withers SG: Engineering of glycosidases and glycosyltransferases. *Current Opinion in Chemical Biology* 2006, **10**(5):509-519.
16. Mackenzie LF, Wang Q, Warren RAJ, Withers SG: Glycosynthases: Mutant glycosidases for oligosaccharide synthesis. *Journal of the American Chemical Society.* 1998, **120**(22):5583-5584.
17. Cobucci-Ponzano B, Strazzulli A, Rossi M, Moracci M: Glycosynthases in biocatalysis. *Advanced Synthesis & Catalysis.* 2011, **353**(13):2284-2300.
18. Wang L-X: Expanding the repertoire of glycosynthases. *Chemistry and Biology.* 2009, **16**(10):1026-1027.
19. Okuyama M, Mori H, Watanabe K, Kimura A, Chiba S: Alpha-glucosidase mutant catalyzes “alpha-glycosynthase”-type reaction. *Bioscience, Biotechnology and Biochemistry.* 2002, **66**(4):928-933.
20. Jahn M, Chen H, Muellegger J, Marles J, Warren RAJ, Withers SG: Thioglycosynthases: Double mutant glycosidases that serve as scaffolds for thioglycoside synthesis. *Chemical Communication.* 2004, **35**(23):no-no.
21. Jahn M, Marles J, Warren RAJ, Withers SG: Thioglycoligases: Mutant glycosidases for thioglycoside synthesis. *Angewandte Chemie International Edition.* 2003, **42**(3):352-354.
22. Kim Y-W, Lovering AL, Chen H, Kantner T, McIntosh LP, Strynadka NCJ, Withers SG: Expanding the thioglycoligase strategy to the synthesis of α-linked thioglycosides allows structural investigation of the parent enzyme/substrate complex. *Journal of the American Chemical Society.* 2006, **128**(7):2202-2203.

23. Müllegger J, Jahn M, Chen H-M, Warren RAJ, Withers SG: Engineering of a thioglycoligase: randomized mutagenesis of the acid–base residue leads to the identification of improved catalysts. *Protein Engineering Design and Selection.* 2005, **18**(1):33-40.

24. Ficko-Blean E, Stubbs KA, Nemirovsky O, Vocadlo DJ, Boraston AB: Structural and mechanistic insight into the basis of mucopolysaccharidosis IIIB. *Proceedings of the National Academy of Sciences.* 2008, **105**(18):6560-6565.

25. Ficko-Blean E, Boraston AB: The Interaction of a Carbohydrate-binding Module from a *Clostridium perfringens* N-Acetyl-β-hexosaminidase with Its Carbohydrate Receptor. *Journal of Biological Chemistry.* 2006, **281**(49):37748-37757.

Chapter 4: Conclusion and future work

4.1 Summary of results achieved

M. tuberculosis utilizes mycothiol (MSH) as the major low molecular weight thiol to protect itself against oxidative stress and so doing ensure its growth and survival [1-5]. This has led to numerous studies on the enzymes involved in the biosynthesis of MSH and MSH-dependent enzymes with the aim to identify potential drug targets [6]. The work described in this thesis focused on MshB (GlcNAc-Ins deacetylase), the third enzyme in the MSH biosynthesis pathway. Currently there are several practical factors that affect the development of inhibitory compounds against MshB, which includes the lack of a suitable high-throughput continuous assay to assay MshB activity, and the poor availability of its natural and alternative substrates. To address these issues two aims were set for this study.

Aim 1: ***Investigation and characterization of known alternative substrates of MshB for use in the development of a new continuous assay of the enzyme***

In the course of this study three chromogenic compounds (GlcNAc-SDNP, GlcNAc-SPNP and GlcNAc-SBn) were used as alternative substrates for MshB for the development of a continuous assay of its activity. Analyses with an FSA-based assay showed that GlcNAc-SBn is the best alternative substrate for MshB of the three substrates, while on the other hand a DNTB-based showed GlcNAc-SDNP to be the best alternative substrate. It was found that the deacetylation of GlcNAc-SDNP by MshB unmasks a free thiol through an S→N intramolecular rearrangement (a so-called Smiles rearrangement) that occurs between the free amino group exposed after the deacetylation reaction and the dinitrophenyl group. The free thiol reacts with DNTB to liberate NTB, the concentration increase of which can be measured continuously by following the absorbance increase at λ_{max} 412nm. Through this enzyme-initiated Smiles rearrangement the first continuous assay for MshB was developed, thereby achieving the first aim of the project.

Aim 2: *Engineering of an α-N-acetylglucosaminidase for biocatalytic preparation of thioglycosides as alternative substrate of MshB*

Three new α-thioglycoligases were generated through site-directed mutagenesis of an α-*N*-acetylglucosaminidase of *Clostridium perfringens* (*Cp*GH89) using SOE PCR, and expression using the Novagen pET™ system to give N-terminally His-tagged proteins that could be purified by IMAC. The generated α-thioglycoligases (*Cp*GH89E483Q, *Cp*GH89E483A and *Cp*GH89E483S) were found to be thermally as stable as the parent *Cp*GH89 glycosidase (Section 3.3.3, Figure 3.8) and were able to promote transglycosylation in the presence of NaN_3. Furthermore, the α-thioglycoligases displayed much reduced hydrolytic activity than the parent *Cp*GH89 glycosidases. However, the *Cp*GH89E483G mutant only promoted hydrolysis even in the presence of NaN_3 (Section 3.3.5, Figure 3.11, R3+ and R3-), and was therefore not found to be suitable for use as a biocatalyst. In summary, three α-thioglycoligases based on *Cp*GH89 were generated, thereby achieving the second aim of the study. These enzymes are good candidates for the enzymatic synthesis of GlcNAc-based α-thioglycosides.

4.2 Future work

4.2.1 Validation of α-thioglycoligases as a biocatalysts for synthesis of thioglycosides

Although the ability of the α-thioglycoligases to promote the transglycosylation step has been confirmed, as well as the reactivity of the glycosyl-enzyme intermediate with the small azide nucleophile, the next important step will be to show that thiol-containing nucleophiles can also be used in this manner. This will validate the biocatalysts' α-thioglycoligase ability, and will allow their implementation in the large-scale production of a range of MSH analogs. Future studies will therefore include the chemical synthesis of suitable thiosugar acceptor molecules, since these are not commercially available, followed by validation of the biocatalyst's α-thioglycoligase ability.

4.2.1.1 Chemical synthesis of thiosugar acceptor molecules

Two different thiosugar acceptor molecules will be synthesised by means of multi-step chemical synthesis based on previously described methods, as outlined in Scheme 4.1 [7-9].

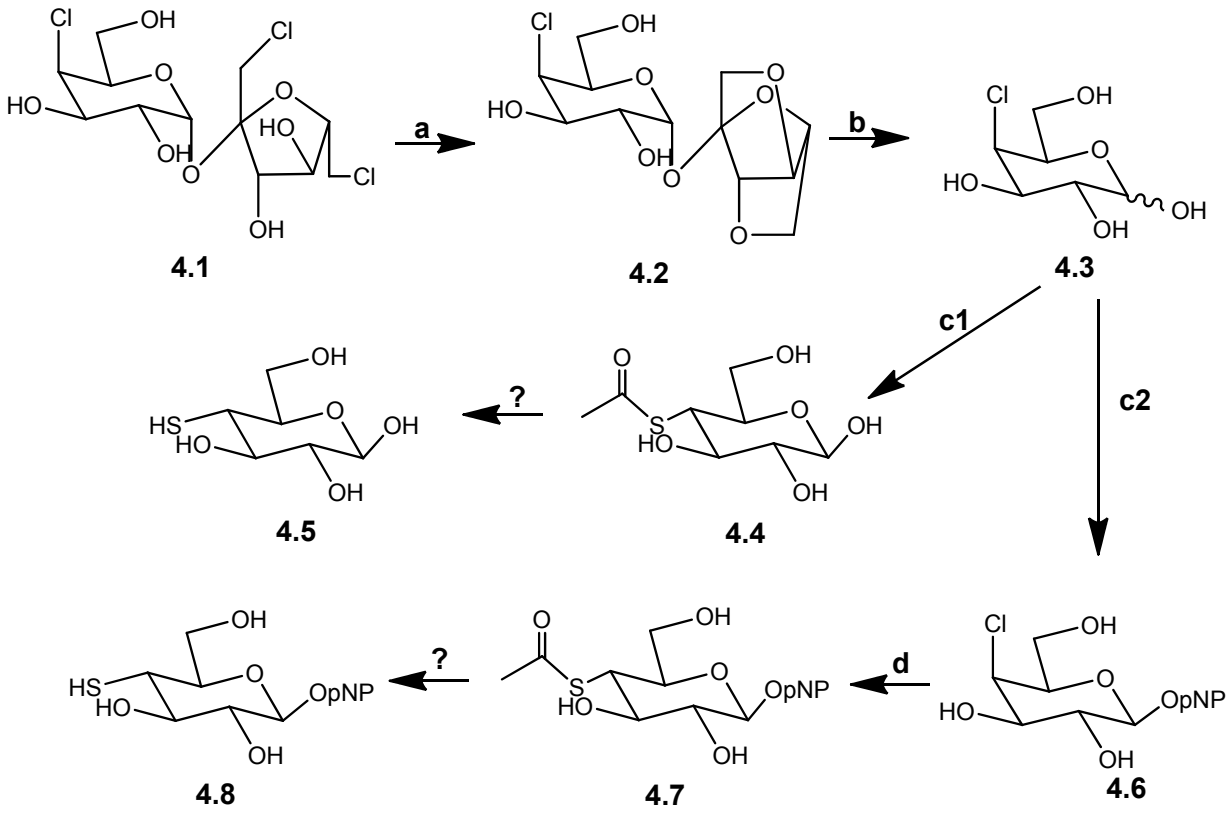
Synthesis of 4-thio-β-D-glucopyranoside (**4.5**): The method of Liu *et al.* [7] will be used to synthesise 1,4:3.6-dianhydro-β-D-fructofuranosyl-4-chloro-4-deoxy-α-D-galactopyranoside (**4.2**), wherein compound **4.1** (sucralose) will be the starting material. Compounds **4.3** (4-chloro-4-deoxy-α-D-galactopyranose) and **4.4** (4-thio-β-D-galactopyranose) will subsequently be synthesised following the method described by Liu *et al.* in 2005, and Chen, *et al*. in 2010, respectively [8, 9].

At present compounds **4.2**, **4.3** and **4.3** have already been prepared using the published methodology. Future work involves the up scaling of the production of compound **4.3**, and finding conditions to perform the thioacetylation reaction to form compound **4.4** followed by the removal of the acetyl group to make the desired compound **4.5** (4-thio-β-D-glucopyranoside).

Synthesis of *p*-nitrophenyl-4-deoxy-4-thio-β-D-glucopyranoside (**4.8**):

The same previously described methods will be used to make compound **4.2** and **4.3**. Compound **4.6** (*p*-nitrophenyl-4-deoxy-4-chloro-β-D-glucopyranoside) will then be synthesised following previously described methods. Compounds **4.7** (*p*-nitrophenyl-4-deoxy-4-thioacetate-β-D-glucopyranoside) and **4.8** will be prepared using the same methodology used to prepare **4.4** and **4.5**.

At present compounds **4.2**, **4.3** and **4.3** have already been produced using the published methodology. Future work involves up scaling the production of compound **4.3,** synthesis of compound **4.6**, and conversion of it to compound **4.7** and compound **4.8** (*p*-nitrophenol-1,4-dithio-β-D-glucopyranoside).



Scheme 4.1: The reagents and condition for the chemical synthesis of the different thiosugar acceptor molecules: (a) 0.5 M of compound **4.1** in 100 mL H_2O, 3.5 M of KOH in 50 mL H_2O, 55^0C, 4h; (b) H_2SO_4 (0.4 mL), 0.25 M of compound **4.2** in 40 mL H_2O, 60^0C; (c1) 5g of compound **4.3**, 105 mL of DMPU (1,3-dimethyl-3,4,5,6-tetra-hydro-2(1H)-pyrimidinone), 8.6 g KSAc, 1.5h, RT; (c2) 0.5g of compound **4.3**, 0.5g of *p*-nitrophenol, 15 mL anhydrous CH_2Cl_2, 1g 4 Å molecular sieves, 50 µl of $BF_3.Et_2O$, 1 mL of Et_3N, 3h, 30^0C, (d) Same as c1, except that the starting material is compound **4.6**.

4.2.1.2 Validation of the biocatalyst's α-thioglycoligase ability and the implementation of the new α-thioglycoligase created for large-scale production of a range of MSH analogs.

The biocatalytic preparation of α-thioglycosides will be done using the approach described in Section 3.2.4, Figure 3.5. The enzymatic reaction products will be purified by preparatory TLC and analyzed by NMR and LC-MS. These synthesized analogs will be evaluated as alternative substrates of MshB using both the discontinuous and continuous assay for MshB described in Chapter 2. In addition, the compounds will also be tested as alternative substrates or potential inhibitors against the other MSH biosynthesis pathway.

4.2.2 Investigation of α-glycosynthases as biocatalysts

As stated in Section 3.2, two options exist for the generation of a biocatalyst that will enable synthesis of natural or alternative substrates for MshB in large quantity. For this study, **Option 2**, namely the use of a thioglycoligase strategy, was investigated as reported in Chapter 3. However, during the course of this study, investigations in the use **Option 1**, which is based on the use of glycosynthases as biocatalysts, were initiated by our collaborator (Dr. M. Moracci, IPB, Naples, Italy). His group generated three potential α-glycosynthases, *Cp*GH89E601G, *Cp*GH89E601A and *Cp*GH89E601S, from the α-*N*-acetylglucosaminidase of *Clostridium perfringens* (*Cp*GH89) by mutating the nucleophilic catalytic active residue Glu601 to Gly, Ala and Ser through site-directed mutagenesis. As part of our on-going collaboration with his group, these mutants were tested for thermal stability. Only *Cp*GH89E601G was found to be thermally as stable as the wild type, while the other two mutants were less stable than the wild type. *Cp*GH89E601G also demonstrated reactivation by an external nucleophile (NaN_3) in reactions analysed by TLC. However, it was also found to promote hydrolysis activity in absence of an external nucleophile (Samples R2+ and R2-, Figure 3.11). This was quite surprising as one would expect hydrolysis activity to occur only in the presence of an external nucleophile that can mimic the activity of the mutated active site nucleophile residue. This indicates that there might be another unidentified nucleophile in the active site, or that the mutation caused a rearrangement in the active site that changed the chemical roles of the active site residues.

Future studies, which will be performed in collaboration with Dr. Moracci, will involve identification of this nucleophile and evaluation of the α-glycosynthase ability of this enzyme for use in the biosynthesis of MSH and its analogs using the approach described in Section 3.2.4.

4.3 Conclusion

The development of the first high-throughput continuous assay of MshB should have a significant positive effect on the on-going drug development studies that target this enzyme, and may allow the identification of a suitable MshB inhibitor with antituberculosis properties. Furthermore, the preparation of thioglycoligase and glycosynthase biocatalysts based on the α-*N*-acetylglucosaminidase from *C.*

perfringens could enhance the availability of the MSH biosynthetic intermediates and perhaps MSH itself, as this will allow their preparation through enzymatic synthesis without the need for the protection and deprotection of the other functional groups as would be required in chemical synthesis. Thus far, the prepared thioglycoligases have been revealed to be good candidates for the preparation of GlcNAc-based thioglycoside molecules which can be used as either alternative substrates or inhibitors of MSH biosynthetic enzyme – especially MshB – and other MSH-dependent enzymes. While this may solve the problem of synthesizing MSH and it analogs, such a strategy will still require the production of the substrate donor and substrate acceptor molecules that are required for the enzymatic synthesis reactions. However, since the synthesis of these molecules are significantly simpler than that of MSH and its biosynthetic intermediates, the biocatalytic preparation method is still an important improvement of the current situation, and should enhance the availability of these molecules for antituberculosis drug development that target MSH biosynthesis.

4.4 References

1. Sareen D, Newton GL, Fahey RC, Buchmeier NA: Mycothiol is essential for growth of *Mycobacterium tuberculosis* Erdman. *Journal of Bacteriology.* 2003, **185**(22):6736-6740.
2. Jothivasan VK, Hamilton CJ: Mycothiol: synthesis, biosynthesis and biological functions of the major low molecular weight thiol in actinomycetes. *Natural Product Reports.* 2008, **25**(6):1091-1117.
3. Rawat M, Av-Gay Y: Mycothiol-dependent proteins in actinomycetes. *FEMS Microbiology Reviews.* 2007, **31**(3):278-292.
4. Dosanjh NS, Rawat M, Chung J-H, Av-Gay Y: Thiol specific oxidative stress response in mycobacteria. *FEMS Microbiology Letters.* 2005, **249**(1):87-94.
5. Fan F, Vetting MW, Frantom PA, Blanchard JS: Structures and mechanisms of the mycothiol biosynthetic enzymes. *Current Opinion in Chemical Biology.* 2009, **13**(4):451-459.
6. Newton GL, Buchmeier N, Fahey RC: Biosynthesis and functions of mycothiol, the unique protective thiol of actinobacteria. *Microbiology and Molecular Biology Reviews* 2008, **72**(3):471-494.
7. Liu F-W, Liu H-M, Ke Y, Zhang J: A facile approach to anhydrogalactosucrose derivatives from chlorinated sucrose. *Carbohydrate Research.* 2004, **339**(16):2651-2656.
8. Liu F-W, Zhang Y-B, Liu H-M, Song X-P: Preparation of α and β anomers of 1,2,3,6-tetra-O-acetyl-4-chloro-4-deoxy-D-galactopyranose based upon anomerization and kinetic acetylation. *Carbohydrate Research.* 2005, **340**(3):489-495.
9. Chen H-M, Withers SG: Syntheses of *p*-nitrophenyl 3- and 4-thio-β-D-glycopyranosides. *Carbohydrate Research.* 2010, **345**(18):2596-2604.

Addendum A

Supplementary Information for the OBC publication (reproduced in Chapter 2)

Supplementary Information

An enzyme-initiated Smiles rearrangement enables the development of an assay of MshB, the GlcNAc-Ins deacetylase of mycothiol biosynthesis

Dirk A. Lamprecht,[a] Ndivhuwo O. Muneri,[a] Hayden Eastwood,[b] Kevin Naidoo,[b] Erick Strauss*[a] and Anwar Jardine*[c]

[a] Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602 South Africa.

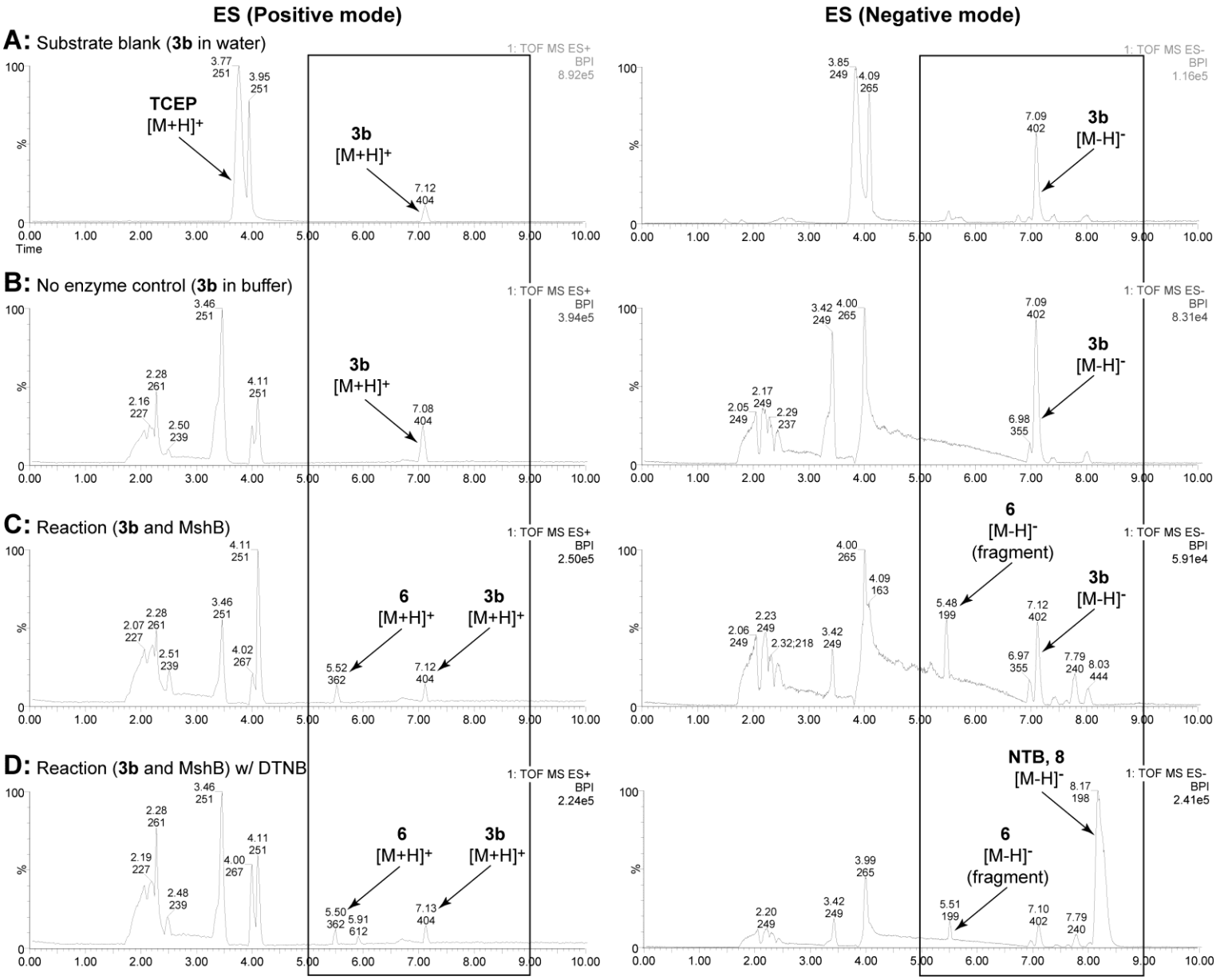
[b] Scientific Computing Research Unit, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa.

[c] Department of Chemistry, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa.

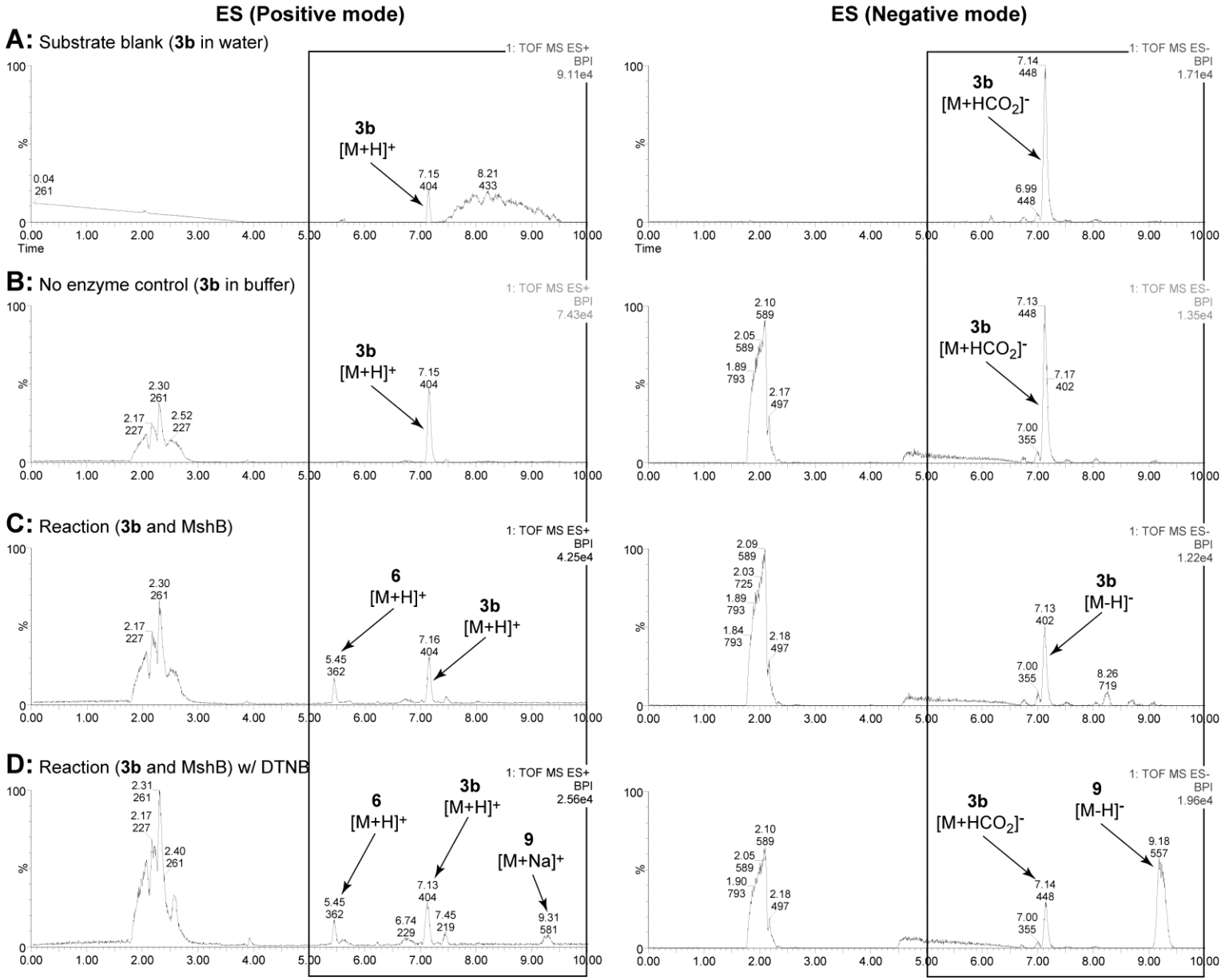
*E-mail: estrauss@sun.ac.za; anwar.jardine@uct.ac.za

LC-HRMS: Base peak ion chromatograms

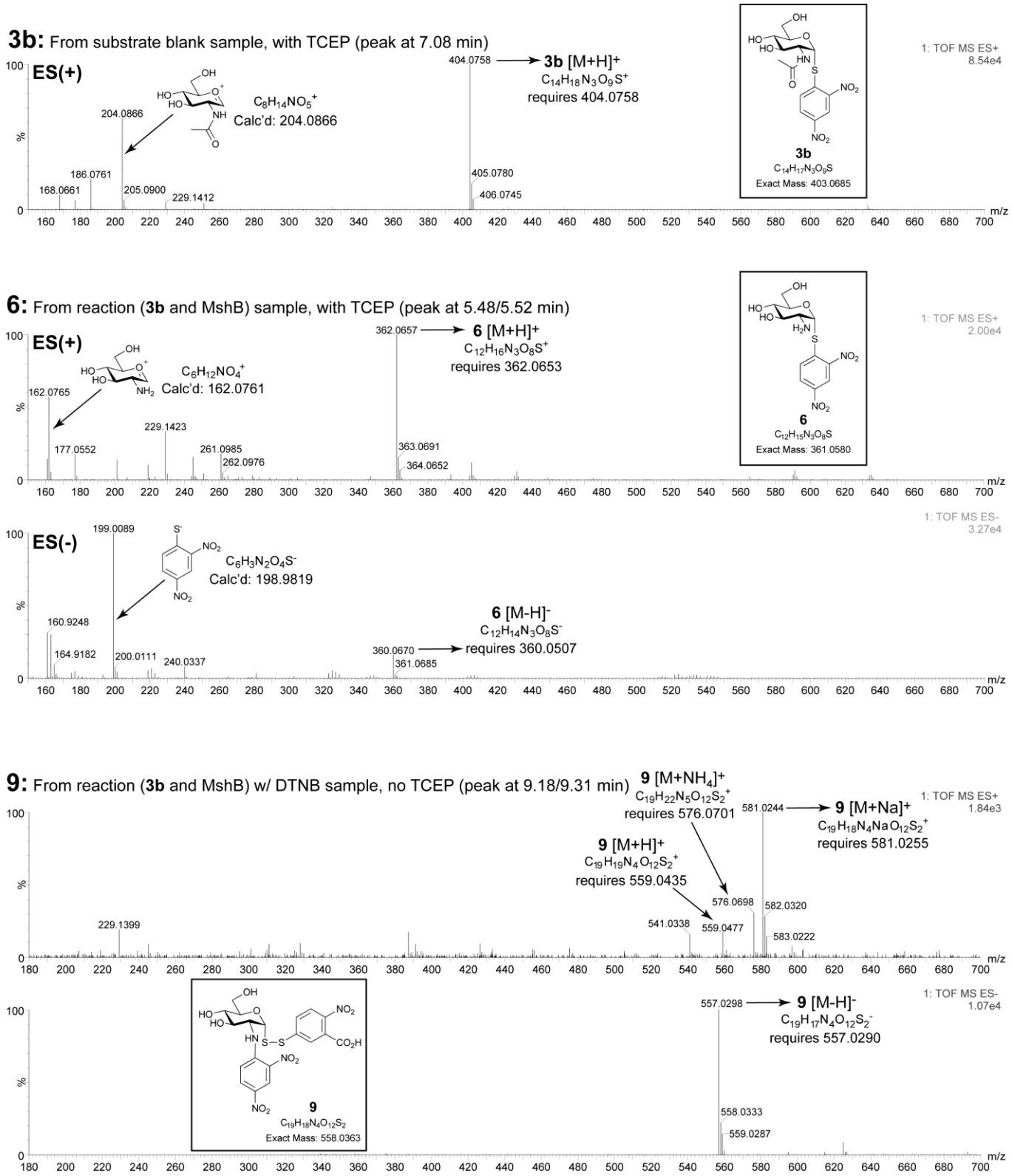
For samples prepared with TCEP added:



For samples prepared without added TCEP:



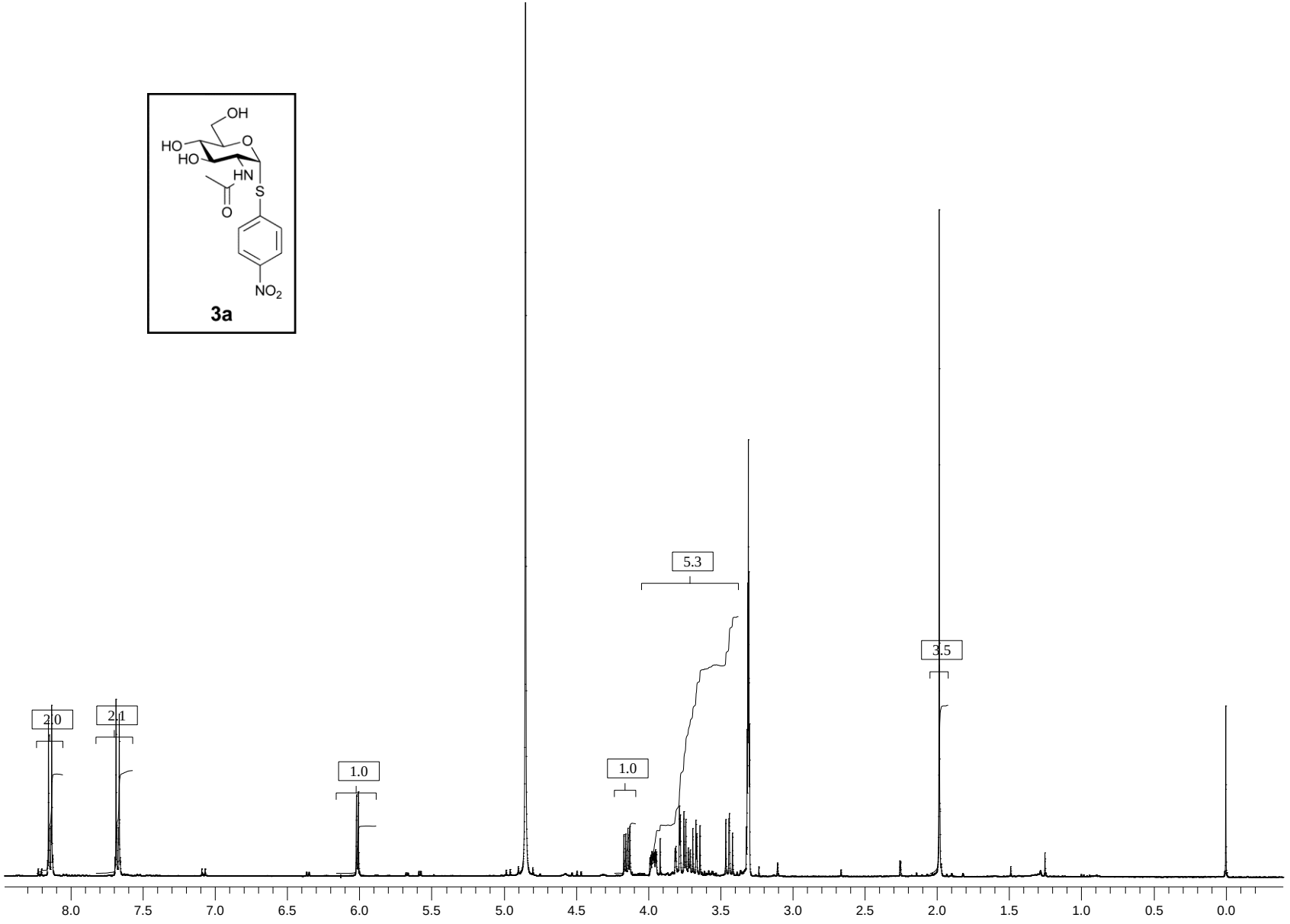
Mass spectra extracted from chromatograms:



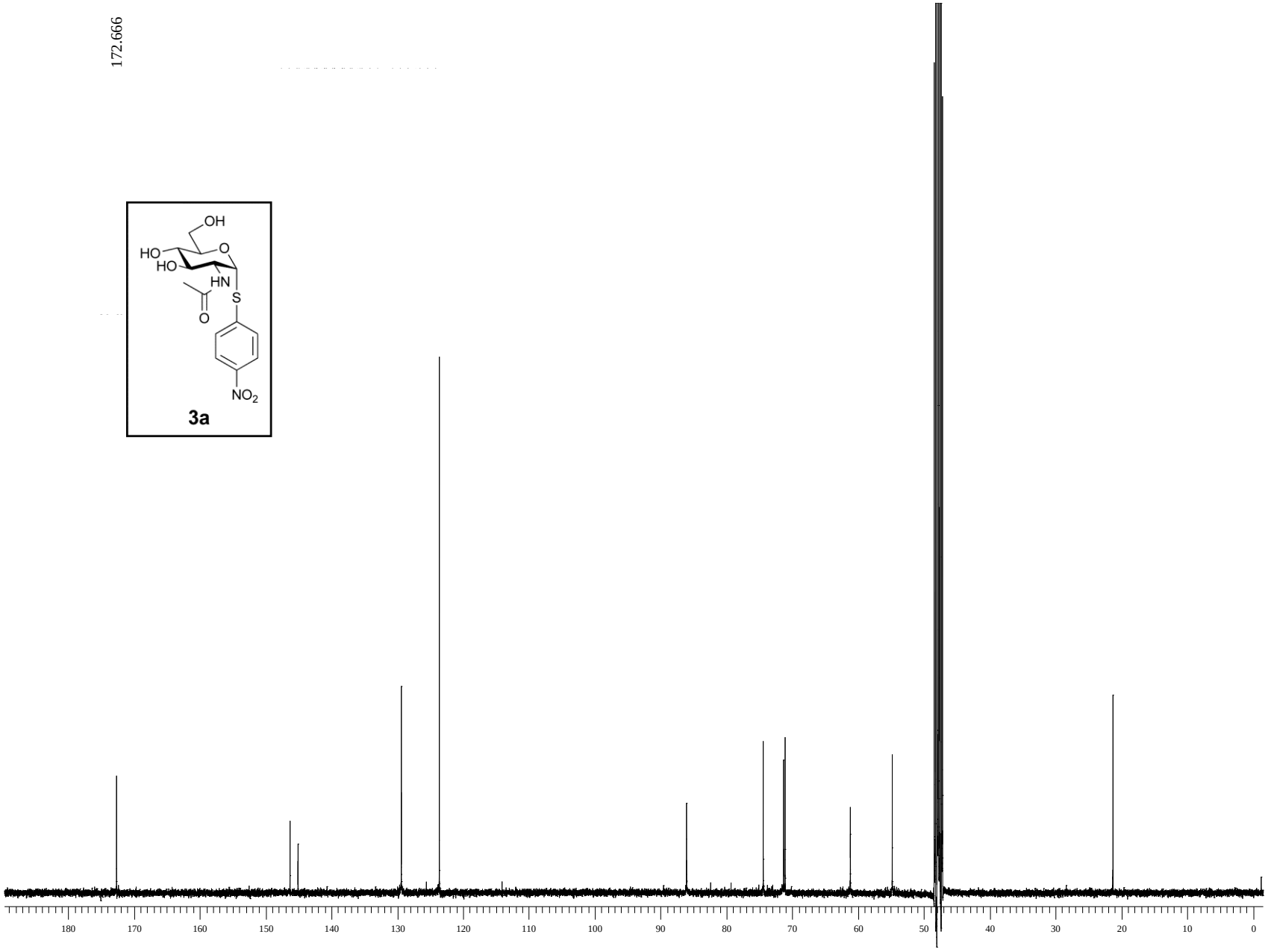
NMR Spectra:

4-Nitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose (3a)

^{1}H NMR (300 MHz, CD_3OD):

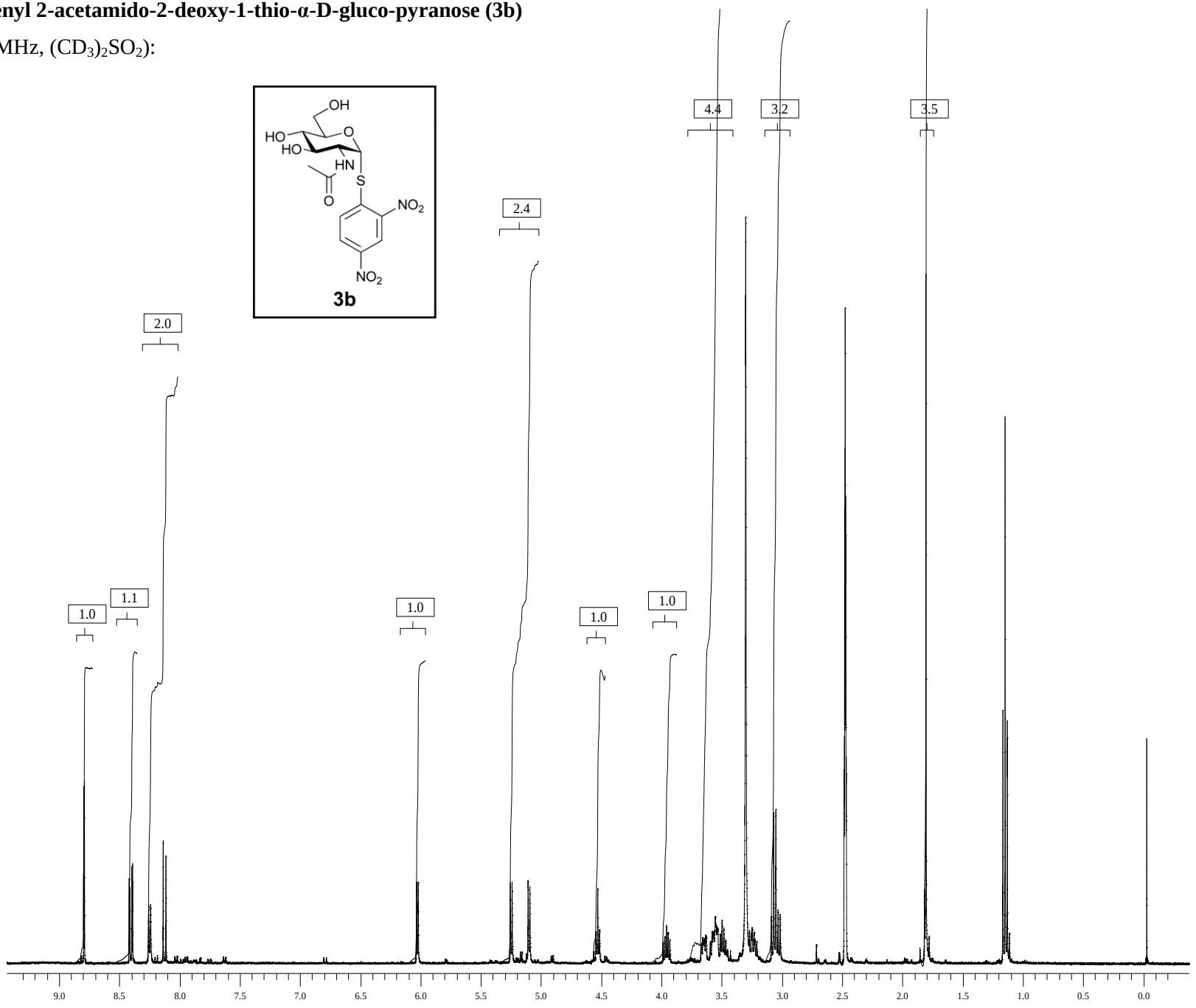


^{13}C NMR (75 MHz, CD_3OD):



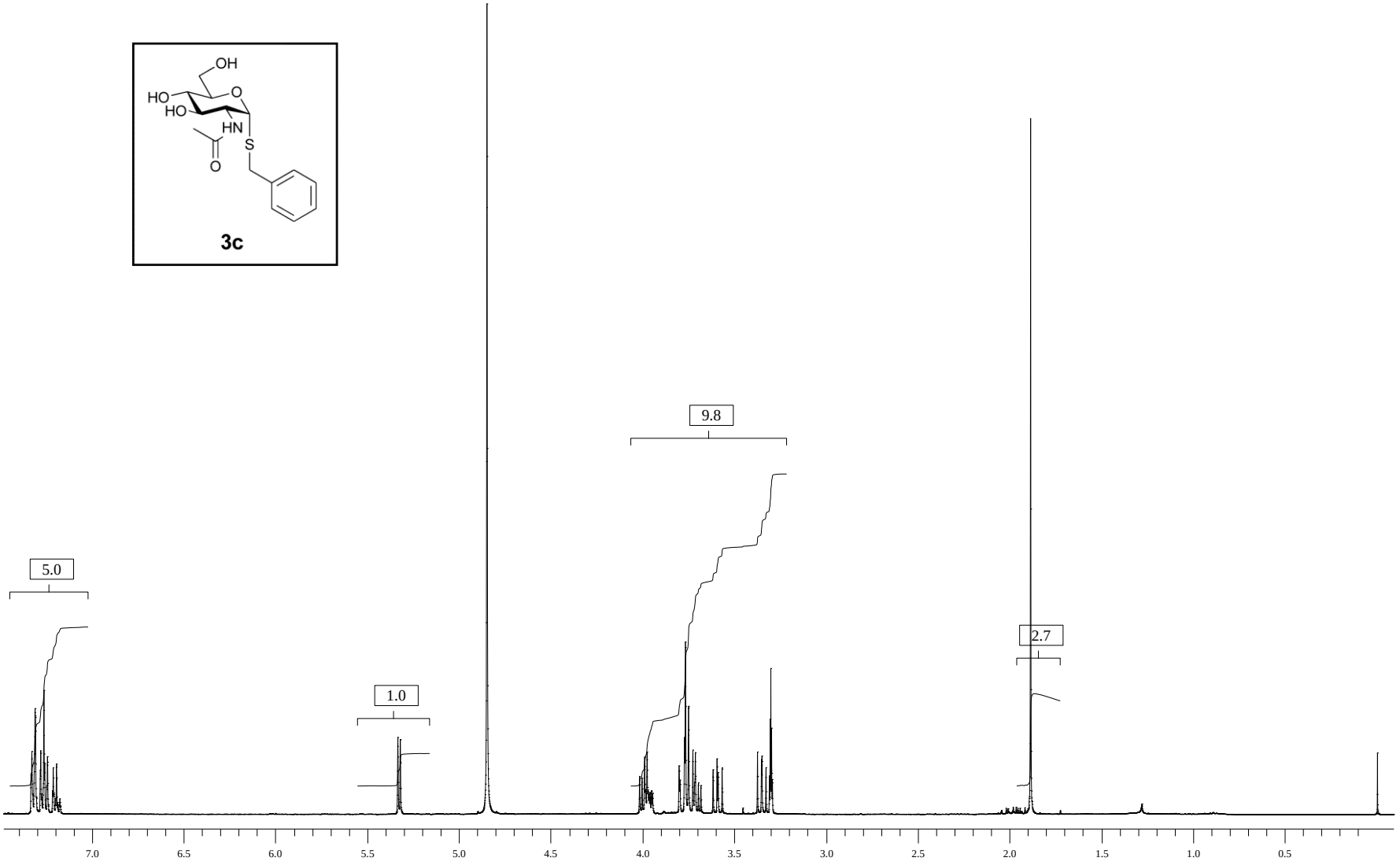
2,4-Dinitrophenyl 2-acetamido-2-deoxy-1-thio-α-D-gluco-pyranose (3b)

^{1}H NMR (300 MHz, $(CD_3)_2SO_2$):



Benzyl 2-acetamido-2-deoxy-1-thio-α-D-glucopyranose (3c)

^{1}H NMR (300 MHz, CD_3OD):



^{13}C NMR (75 MHz, CD_3OD):