Mechanistic characterisation and inhibitor identification of *Mycobacterium tuberculosis* bifunctional *N*-acetyltransferase/uridylyltransferase GlmU

Peter D. Craggs

University College London
and
The Francis Crick Institute
PhD Supervisor: Dr Luiz Pedro Sório de Carvalho

A thesis submitted for the degree of
Doctor of Philosophy
University College London
May 2019

Declaration

I Peter David Craggs confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The mycobacterial cell envelope is a defining feature of the bacteria, primarily due to its highly-ordered, relatively impermeable nature that is likely one of the key attributes that has contributed to the success of this pathogenic over the last thousand years. Peptidoglycan is a unique and essential structural element that provides much of the strength and rigidity of the mycobacterial cell envelope. Most of the enzymes involved in the biosynthetic pathway of peptidoglycan have been shown to be essential for *Mycobacterium tuberculosis* growth.

Mycobacterium tuberculosis GlmU is an essential bifunctional *N*-acetyltransferase, uridylyltransferase enzyme involved in the formation of uridine-diphosphate *N*-acetylglucosamine, which is the universal donor of *N*-acetylglucosamine for both peptidoglycan and lipopolysaccharide biosynthesis. This enzyme catalyses acetylation of glucosamine 1-phosphate, followed by uridylylation of *N*-acetylglucosamine 1-phosphate.

Detailed characterisation of the kinetic mechanism ascertained that acetyl transfer progresses by the formation of a ternary complex, with acetyl coenzyme A binding preceding glucosamine 1-phosphate and coenzyme A the last product to dissociate. A novel ternary complex crystal structure, with glucose 1-phosphate and acetyl-coenzyme A, identified a candidate general base involved in the deprotonation of glucosamine 1-phosphate, as well as other important active site residues for substrate binding and catalysis.

pH-rate studies and site-directed mutagenesis led to assignment of Histidine 374 as the catalytic general base. Solvent kinetic isotope effect experiments and pH-rate studies identified that acetyl transfer is partially rate-limiting.

Small molecule screening led to the identification of novel inhibitors of GlmU catalysed acetyl transfer reaction. Inhibitors identified from the GSK TB set, were confirmed as GlmU interacting compounds and shown to inhibit the growth of *M. tuberculosis*. These results show that knowledge of the kinetic and catalytic mechanism enabled optimisation of a thorough screening approach that identified novel inhibitors that demonstrated that GlmU acetyltransferase activity is essential for *M. tuberculosis* growth.

Impact Statement

The work presented in this thesis initially focusses on characterising the kinetic and chemical mechanism of the left-handed beta helix (L*β*H) domain containing acetyltransferase enzyme, GlmU. Further studies focussed on identifying novel small molecule inhibitors of the GlmU acetyltransferase activity and the subsequent confirmation that these compounds can inhibit the growth of *Mycobacterium tuberculosis*. In addition to catalysing acetyl transfer from acetyl-coenzyme A (Ac-CoA) to glucosamine 1-phosphate (GlcN-1P), GlmU is also a pyrophosphorylase, which catalyses uridylyl transfer from uridine-triphosphate (UTP) to *N*-acetylglucosamine 1-phosphate (GlcNAc-1P), a product of the acetyltransferase activity. The formation of uridine-diphosphate *N*-acetylglucosamine (UDP-GlcNAc) is essential in most bacteria as it is the universal donor of *N*-acetylglucosamine (GlcNAc) for both peptidoglycan and lipopolysaccharide biosynthesis. These biosynthetic pathways are critical for the formation of the cell envelope of *Mycobacterium tuberculosis* (*M. tuberculosis*), the causative agent of tuberculosis (TB). Genetic target validation experiments, such as Transposon Hybridisation (TraSH) studies, identified several of the enzymes involved in both the peptidoglycan and lipopolysaccharide biosynthetic pathways as essential for *M. tuberculosis* growth, including GlmU.

The global impact of TB is still hugely significant, in 2017, the disease was the ninth leading cause of death worldwide, with an estimated 1.3 million fatalities. The emergence and spread of antibiotic resistant TB is an increasing global health threat. For instance, in 2017, there were 558,000 cases of multidrug-resistant TB (MDR-TB), with 8.9% of these incidences estimated to be extensively drug-resistant TB (XDR-TB). To exacerbate this issue, only two new TB drugs have been developed in the last 40 years and global spending on TB drug discovery and development is still well below the estimated required levels. The magnitude of this global health threat underscores the urgent need to discover new antimycobacterial agents that have novel mechanisms of action (MOA). To this end, it is critical to understand how genetically validated, essential *M. tuberculosis* enzymes, such as GlmU, which could be attractive drug targets, function.

There were many facets of the GlmU acetyltransferase mechanism unknown, or kinetically untested, at the outset of this project, which are important to define and understand prior to progressing with drug discovery studies. The aim of this study is to thoroughly characterise the kinetic and chemical mechanisms of the acetyltransferase activity of GlmU and to utilise this information to develop a robust, sensitive small molecule screening strategy that could be utilised to screen both synthetic compound libraries, as well as initiating knowledge-based inhibitor design. Additionally, this project has highlighted the combination of drug discovery strategies, utilising the results of *M. tuberculosis* phenotypic HTS studies, as effective compound collections for the identification of inhibitors by a target-based approach. The final aims of this project are to both demonstrate the chemical tractability of the acetyltransferase activity of GlmU, while ultimately, using novel inhibitors to prove the essentiality of this enzyme for *M. tuberculosis* growth. Demonstration of the chemical tractability, or druggability, of GlmU acetyltransferase activity has provided evidence that it is possible to identify and confirm drug-like inhibitors. Identification of dual inhibitors of both GlmU catalysed acetyl transfer and *M. tuberculosis* growth has confirmed the essentiality of the enzyme.

This approach of thorough enzyme characterisation, in combination with hit identification utilising phenotypically validated compounds, could be systematically applied to other genetically validated, essential *M. tuberculosis* enzymes.

Acknowledgement

I would like to start by acknowledging my PhD supervisor Luiz Pedro Carvalho for allowing me the opportunity to embark on this project. Luiz has been a supportive and enthusiastic supervisor and has taken the time to help interpret and discuss data, in addition to guiding the next steps. Luiz has taught me many core enzymology techniques and has encouraged me to think small, by concentrating on the details of each experiment and reinforcing the importance of thoroughly interrogating data. Luiz has also supported my attendance of the Gordon Research Conference on Enzymology in the USA for three consecutive years, which has allowed me to learn from discipline experts from around the world.

Argyrides (Archie) Argyrou was instrumental in the genesis of this project, providing the initial contact with Luiz, as well as extensively contributing to the initial proposal. Archie is a hugely knowledgeable enzymologist and his guidance and advice has been invaluable throughout this project.

This project would not have been possible without the initial support from David Powell, who, as my department head at GSK was responsible for providing the financial assistance required. Dave also showed trust and belief in my abilities and has remained a supportive influence throughout the project. Melanie Leveridge, who succeeded Dave as department head, has continued to provide invaluable support, which has enabled the completion of this project.

I'd like to acknowledge and thank Stephane Mouilleron for carrying out all GlmU X-ray crystallography experiments and taking the time to discuss the data and contribute to a publication.

I'd like to thank Cesira de Chiara for carrying out the GlcN-1P pH titration ^{1}H NMR Spectroscopy experiments and allowing me to be involved in both the practical and analytical aspects of this study.

UDP-GlcN was provided by Rob Field and Martin Rejzek, which enabled the inclusion of this reagent in a panel of analogues for substrate specificity experiments. This reagent is not commercially available, and I'd like to take this opportunity to thank both Rob and Martin for their help.

The GlcN-1P inhibitor analogues were kindly contributed by Ben Whitehurst and Rob Young, who took the time to design and make these novel compounds. The

synthesis and purification of the GlcN-1P inhibitor analogues were one of the subjects of Ben's PhD project. Collaboratively working with both Ben and Rob was a productive and rewarding experience. I'd also like to thank Rob for his continued chemistry guidance and tuition throughout this project.

Ryan Bingham, my line manager at GSK, throughout most of this project, was largely responsible for developing me as a scientist prior to the initiation of this PhD. Ryan has helped balance my GSK and PhD workload, while continually supporting and encouraging me. Ryan also kindly read this thesis and provided numerous suggestions and edits. Stephen Ashman acted as my independent, industry PhD supervisor and provided support and advice throughout this project. Steve also achieved his PhD while working at GSK, so his perspective on balancing workload has been invaluable.

I'd like to thank my thesis committee of Edgar Deu and Martin Webb for great discussions and the helpful suggestions made during these meetings.

Great technical contributions, practical help and knowledgeable guidance were provided by the following GSK colleagues; Michelle Pemberton (RF-MS), Gareth Cooper (SDM, HT protein expression and purification), Emma Jones (protein purification), Angela Bridges (SDM and protein expression), Ross Hryczanek (chemistry), Gemma White (chemistry) and Anshu Pal (protein purification).

On a personal level, I will be evermore grateful to my eternally patient wife, Sian, who has shown unfaltering belief in me throughout the duration of this PhD project. Sian has also supported me in many ways and without her I would have not been able to complete this PhD. I'd like to thank my mother-in-law, Ann, who has provided invaluable support to both Sian and I throughout this project. I'd like to thank my Mum and Dad, Margaret and David, as well as my brother, Mark, for their continuing pride, encouragement and belief in my abilities. I'd like to thank my wider family and friends for their constant backing throughout this project. Finally, I'd like to dedicate this PhD project and thesis to a very patient four-year-old girl, my daughter, Erin Eliza Craggs, who has never known a world without her Dad working all the time! I can't wait to readdress the balance and spend more time with her in the future.

Table of Contents

Table of figures

List of tables

List of schemes

Abbreviations

4-TP	-	4-thiopyridone
AAc-CoA	-	Acetoacteyl-CoA
A	-	Alanine
Ac-CoA	-	Acetyl-CoA
ACN	-	Acetonitrile
ADP	-	Adenosine 5’-diphosphate
Ala	-	Alanine
AMP	-	Adenosine 5’-monophosphate
Arg	-	Arginine
ART	-	Anti-retroviral therapy
Asn	-	Asparagine
ATP	-	Adenosine 5’-triphosphate
AZ	-	AstraZeneca
Bu-CoA	-	Butyryl-CoA
Ca^{2+}	-	Calcium
$CaCl_2$	-	Calcium chloride
CD	-	Circular dichroism
CDP	-	Cytidine 5’-diphosphate
CFR	-	Case fatality ratio
CFU	-	Colony-forming unit
CoA-SH	-	Coenzyme A
Co^{2+}	-	Cobalt
$CoCl_2$	-	Cobalt chloride
COSY	-	homonuclear correlation spectroscopy
CHAPS	-	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CMP	-	Cytidine 5’-monophosphate
Cro-CoA	-	Crotonyl-CoA
CTP	-	Cytidine 5’-triphosphate
D_2O	-	Deuterium oxide
Dethio-CoA	-	CoA without free thiol group
DMSO	-	Dimethyl sulphoxide
ΔT_m	-	Change in melting temperature

DS-TB - Drug-susceptible tuberculosis
DSF - Differential scanning fluorimetry
DTP - 4,4′-Dithiodipyridine
E. coli - *Escherichia coli*
EDTA - Ethylenediaminetetraacetic acid
ESI - Electrospray ionisation
Eth-CoA - Ethyl-CoA
FAC - Final assay concentration
FDA - US Food and Drug Administration
Gal-1P - Galactose 1-phosphate
GalN-1P - *N*-acetylgalactosamine 1-phosphate
GalN-6P - *N*-acetylgalactosamine 6-phosphate
GDP - Guanosine 5'-diphosphate
Glc-1P - Glucose 1-phosphate
Glc-6P - Glucose 6-phosphate
GlcN - Glucosamine
GlcN-1P - Glucosamine 1-phosphate
GlcN-6P - Glucosamine 6-phosphate
GlcNAc - *N*-acetylglucosamine
GlcNAc-1P - *N*-acetylglucosamine 1-phosphate
GlmU - GlcN-1P acetyltransferase uridylyltransferase
glmU - Gene encoding GlmU
GlmM - Phosphoglucosamine mutase
GlmS - Glucosamine 6-phosphate synthase
GMP - Guanosine 5'-monophosphate
GSK - GlaxoSmithKline
GTP - Guanosine 5'-triphosphate
H_2O - Hydrogen oxide, water
HEPES - N-(2-Hydroxyethyl)piperazine-N-(2-ethanesulfonic acid)
H - Histidine
H^+ - Hydrogen ion
HCl - Hydrogen chloride
His - Histidine
HRP - Horseradish peroxidase

HSQC - Heteronuclear single quantum correlation
HTS - High throughput screen
IC_{50} - Half maximal inhibitory concentration
IsoBu-CoA - Isobutyryl-CoA
K - Lysine
K - Michaelis constant
K_a - Acidic pK_a constants
K_{act} - Activation constant
K_b - Basic pK_a constants
k_{cat} - Turnover number
K_d - Dissociation constant
kDa - Kilo Daltons
K_i - Dissociation constant for the enzyme-inhibitor complex
K_{ii} - Competitive inhibition constant, or intercept inhibition constant
K_{is} - Uncompetitive inhibition constant, or slope inhibition constant
K_m - Michaelis constant
LβH - Left-handed beta helix
LTBI - Latent-tuberculosis infection
Lys - Lysine
Man-1P - Mannose 1-phosphate
Man-6P - Mannose 6-phosphate
ManN - Mannosamine
ManNAc - *N*-acetylmannosamine
ManN-1P - Mannosamine 1-phosphate
MES - 4-Morpholineethanesulfonic acid
MDR-TB - Multidrug-resistant tuberculosis
Mg^{2+} - Magnesium
$MgCl_2$ - Magnesium chloride
Mlo-CoA - Malonyl-CoA
Mn^{2+} - Manganese
$MnCl_2$ - Manganese chloride
MIC - Minimum inhibitory concentration
MS - Mass spectrometry
MST - MicroScale Thermophoresis

Mtb - *Mycobacterium tuberculosis*

n - Number of replicates

N - Asparagine

NaCl - Sodium chloride

NanoDSF - Capillary based DSF

NaOH - Sodium hydroxide

n_H - Hill coefficient

NEM - *N*-Ethylmaleimide

Ni^{2+} - Nickel

$NiCl_2$ - Nickel chloride

Ni-NTA - Nickel-nitrilotriacetic acid

NIMR - National Institute for Medical Research

NMR - Nuclear magnetic resonance

NOESY - Nuclear Overhauser Effect Spectroscopy

PBP - Phosphate binding protein

PDB - Protein Data Bank

PEG - Polyethylene glycol

PG - Peptidoglycan

pK_a - Acid association constant

pK_d - negative $\log_{10} K_d$

pK_i, - negative $\log_{10} K_i$

pH - decimal logarithm of the reciprocal of the hydrogen ion activity

Pi - Inorganic phosphate

pIC_{50} - negative $\log_{10} IC_{50}$

PNP - Purine nucleoside phosphorylase

PPi - Inorganic pyrophosphate

PPase - Inorganic pyrophosphatase

Pro-CoA - *n*-Propionyl-CoA

OD - Optical density

R - Arginine

r^2 - Coefficient of determination

RF-MS - RapidFire MS

RFU - Relative fluorescence units

SAR - Structure activity relationship

SD - Standard deviation
SDM - Site directed mutagenesis
SDS-PAGE - Sodium dodecylsulfate – polyacrylamide gel electrophoresis
SE - Standard error of the mean
SKIE - Solvent kinetic isotope effect
SNP - Single nucleotide polymorphism
SPR - Surface plasmon resonance
STD - Saturation transfer difference
Suc-CoA - Succinyl-CoA
T_m - Melting temperature
T_m^{app} - Apparent melting temperature
TB - Tuberculosis
TTP - Thymidine 5'-triphosphate
Tyr - Tyrosine
UDP - Uridine 5'-diphosphate
UDP-GlcNAc - UDP-*N*-acetylglucosamine
UDP-GlcN - UDP-glucosamine
UMP - Uridine 5'-monophosphate
UTP - Uridine 5'-triphosphate
v - k_{cat}
V_0 - Initial velocity
V_{max} - Maximum velocity
WT - Wild type
XDR-TB - Extremely drug-resistant tuberculosis
XOD - Xanthine oxidase
Y - Tyrosine
Zn^{2+} - Zinc
$Zn(OAc)_2$ - Zinc acetate
$ZnCl_2$ - Zinc chloride

Chapter 1. Introduction

1.1 Tuberculosis and *Mycobacterium tuberculosis*

1.1.1 Tuberculosis

Tuberculosis (TB) is an infectious disease that is predominantly caused by the bacillus *Mycobacterium tuberculosis* that was first isolated in pure culture by Robert Koch in Berlin in 1882[1]. TB is a disease found in many mammals and the human condition is mainly caused by *M. tuberculosis*, but can also be caused by *M. bovis*, *M. africanum*, and *M. microti*, which taken together are referred to as the *M. tuberculosis* complex (MTBC)[2-6]. *M. tuberculosis* can be found throughout the world and is not thought to cause TB in other animals, although there are some recent reports to the contrary, with humans thought to be the sole, natural reservoir[7]. TB is an ancient disease, with incidences evident in the mummified remains of humans in Egypt and there is increasing evidence that the migration of humans from Africa led to the spread of MTBC, leading to a theory that the two species co-evolved[8-10]. TB, also previously known as consumption or the great white plague, has long blighted humankind but the advent of the industrial revolution with the associated increase in living standards in developing countries, along with increased knowledge of the disease, leading to isolation of infectious diseases all contributed to a decline in the reported mortality rates. The development of the *M. bovis* bacillus Calmette-Guerin (BCG) vaccine, along with the first anti-TB drugs, raised hopes of TB prevention and simultaneously ended the isolation of TB patients in sanatoriums. Unfortunately, the use of TB vaccines and drugs in resource-poor, developing countries, was poor, coupled with the neglect of TB control worldwide in the late-1980s led to a resurgence of TB in the 1990s. Several major, global disruptions, such as the end of the Cold War, the dissolution of the former Soviet Union and the growth of the HIV/AIDS pandemic, all contributed to TB becoming a leading cause of mortality in the developing world, most specifically in sub-Saharan Africa[11].

The use of anti-TB drug combinations, along with a new healthcare strategy, known as directly observed therapy (DOTS), which involved the direct supervision of drug intake by healthcare professionals, contributed to a regaining of control of TB treatment, along with re-establishment of global TB surveillance by the World Health

Organisation (WHO). These efforts have led to a decrease in both global TB incidences and deaths since the late 1990s but there is no escaping the fact that this disease is still a major human health issue. The increasing reports of drug resistant TB, set amongst dire warnings of wider antibiotic drug resistance, highlight that action is urgently required to address these problems and find new therapeutics to prevent the persistence of TB.

1.1.2 Tuberculosis, the disease and the global impact

Human tuberculosis (TB) is one of the most persistent and devastating infectious human diseases of all-time, with one in five adults in Europe and North America killed by the infection between the seventeenth and nineteenth centuries[12,13]. The WHO recently reported that in 2017, TB was the ninth leading cause of death worldwide, with an estimated 10.0 million new cases of active TB, globally[14]. In 2017, there were also 558,000 incidences of multidrug-resistant TB (MDR-TB), which is classified as a *M. tuberculosis* infection that is resistant to treatment with either of the first line TB drugs, isoniazid and rifampicin[15].

Most TB infections, up to 90%, are symptom-less, which are commonly known as latent TB infections (LTBI), that is defined as an infection by *M. tuberculosis* that is contained by the immune system of the host, controlling the replication of the bacteria, and thereby avoiding tissue damage and clinical symptoms. The WHO 2018 Global Tuberculosis Report, estimated 1.7 billion people, which is 23% of the world's population, had a latent TB infection and of these individuals, approximately one in ten, would progress to active TB infection[14].

The TB mortality rate is reducing by 3% per year and the overall reduction during the period from 2000 to 2017 is 42%. Worldwide, the incidence of new cases of TB is reducing by 2% per year, which demonstrates that slow progress is being made in both the management and treatment of the disease. However, to achieve the milestones set out in the WHO End TB Strategy, the number of TB incidences need to reduce by 4 – 5% per year by 2020[16]. Additionally, TB is still the leading cause of death among people suffering from HIV infection, accounting for approximately 35% of all AIDS deaths globally. Two thirds of the global burden of TB is shared by eight countries: India (27%), China (9%), Indonesia (8%), the Philippines (6%), Pakistan (5%), Nigeria (4%), Bangladesh (4%) and South Africa (3%), highlighting

that the disease is mainly found in developing countries where extreme poverty is more prevalent. Finally, drug resistant TB is still a major global public health crisis and the total number of people with MDR-TB increased by 68,000 cases from 490,000 in 2017 to 558,000 in 2018. Worryingly, by the end of 2017, 127 WHO member states had reported confirmed cases of extensively drug-resistant TB (XDR-TB), which is defined as a *M. tuberculosis* infection that is resistant to isoniazid and rifampicin, as well as any fluoroquinolone and at least one of the three second-line injectable drugs (primarily amikacin, or, kanamycin and capreomycin)[15]. In 2017, 230,000 TB deaths were attributed to either MDR or rifampicin resistant (RR)-TB, with MDR-TB being responsible for 82% of these cases. A large proportion of the drug resistant TB cases appeared in India, China and the Russian Federation, accounting for 45% of the total number of reported MDR-TB cases[14].

As previously mentioned, in 2015, the WHO set up the End TB Strategy, to tackle the global TB epidemic, with targets of reducing the number of deaths from TB by 90 and 95%, as well as reducing the incidence rate of TB by 80% and 90%, by 2030 and 2035, respectively, when compared with 2015[16]. To achieve these goals, additional, intermediate milestones to reduce the incidence of TB and the number deaths reported in 2015 by 20% and 35%, by 2020, as well as 50% and 75% reductions by 2025. The proportion of people with active TB who die from the disease, known as the case fatality ratio (CFR), needs to be reduced to 10% globally by 2020. However, the annual reduction in TB incidence rate from 2015 has only reduced by between 1.5 to 2% per year, and the CFR in 2018 was reduced to 16%, revealing that considerable effort needs to be made to meet the End TB Strategies milestones for 2020. From the publication of the End TB Strategy in 2015 to 2017, the countries that had two thirds of the global incidences for TB has risen from six to eight countries, with the addition of Bangladesh and the Philippines. The rate of progress in controlling and preventing the spread of TB in these countries is going to have a major impact on the WHO End TB milestones and the increase in the number of the countries will only serve to complicate the situation and make achieving these goals yet more challenging.

1.1.3 Tuberculosis and the host response

TB is spread between host and potential hosts by the transfer of *M. tuberculosis* by the inhalation of droplets (aerosols) containing a small number of bacilli[17]. TB infection primarily affects the lungs, which is known as pulmonary TB, but can also infect many other parts of the body. Non-pulmonary, or extrapulmonary, TB infections are most commonly found in the lymph nodes, pleura, gastrointestinal tract, bone, CNS and genitourinary system. Manifestations of extrapulmonary TB infection vary on the site of disease, making diagnosis challenging, as in many cases the symptoms are close to, or mimic, other diseases[18]. The pathogenesis of pulmonary TB, on the other hand, has been exhaustively studied and is well understood. Typically, pulmonary TB infection begins when *M. tuberculosis* reaches the pulmonary alveoli, where they invade and replicate within endosomes of alveolar macrophages[19]. Phagocytic cells, such as macrophages, identify the invading bacteria as foreign, mediated through an array of different receptor molecules, including dectin-1, the complement receptor 3 and Toll-like receptors[20]. The phagocytic cells then attempt to eliminate the *M. tuberculosis* cells by phagocytosis, which involves the macrophage engulfing the bacteria to form phagosomes that then subsequently interact with early and late endocytic organelles. This process is known as phagosome maturation, which involves dynamic interactions with different membrane-bound structures that acts to change the membrane and luminal composition of the phagosome[21]. The primary aim of phagosome maturation is to restrict the growth of the internalised pathogen. As the phagosome matures then, typically, it fuses with lysosomes leading to acidification and degradation of the pathogen. However, there is significant evidence to suggest that phagosomes-containing *M. tuberculosis* have deficient luminal acidification, which has subsequently been demonstrated to be due to arrested phagosome maturation induced by the pathogen[22]. Virulent strains of *M. tuberculosis* can escape the phagosome by damaging the membrane, however the mechanism for this has yet to be fully elucidated. Subsequently, the damaged phagosomes are known to be rapidly recognised by the selective autophagy machinery, a major intracellular degradation mechanism involved in recycling of long-lived proteins and cytoplasmic organelles[23]. Phagosome membrane disruption allows *M. tuberculosis* to freely localise in the cytosol, however, exactly how the pathogen enters the cytosol, how

long it remains there, and the mechanism of host cell recognition remains to be fully characterised[22]. *M. tuberculosis* then replicate within the infected macrophages, before a proportion of the host immune cells become necrotic, triggered by the invading mycobacteria, which leads to macrophage cell death and release of *M. tuberculosis* into the extracellular milieu[24,25].

The hallmark of TB infection is the formation of the granuloma, which comprises infected macrophages, along with lymphocytes and other host immune cells, and forms in response to persistent stimuli. The formation of the granulomas is crucial for controlling and containing the TB infection but, as previously discussed, may also provide an environment for *M. tuberculosis* proliferation and dissemination[26]. The TB infection can progress from the granulomas and the mycobacteria can spread throughout the lung, or in the case of severe forms of TB, gain entry to the blood stream and can spread throughout the body, leading to multiple foci of infection.

1.1.4 The history of Tuberculosis and *Mycobacterium tuberculosis*

M. tuberculosis is a member of the phylum Actinobacteria, in the family *Mycobacteriaceae*, which is comprised of a single genus, *Mycobacterium.* There are over a 170 species of *mycobacteria*, all of which are obligate aerobic, non-motile and non-spore forming organisms. Classification using the Gram-staining procedure is not possible as the stain is not well taken up, however, using the Ziehl-Neelsen stain, *Mycobacterium* can be classified as acid-fast bacteria. *Mycobacterium* have a slightly curved or straight rod-like morphology, which is generally between 0.2 and 0.6 µm wide and between 1.0 and 10 µm long**.** A defining characteristic of *mycobacteria* are their unique, highly ordered and complex cell envelope structures, which are rich in both sugars and very long lipids, known as mycolic acids[27]. The cell wall core consists of three primary components; an inner layer of cross-linked polymer of peptidoglycan, which is covalently attached to middle section of highly branched arabinogalactan polysaccharide that is in turn esterified with an outer coating of long-chain mycolic acids[28]. The cell wall is covered in two additional layers; an outer membrane, comprised of non-covalently linked glycophospholipids and inert waxes, and a loosely attached capsule, made up of polysaccharides, proteins and a minor amount of lipids[29]. The origins of the atypical *mycobacteria* cell wall are unclear, it could be a relic, sharing a common origin with the double

membrane present in typical Gram-negative species, or it could be the result of a more recent adaptation[30].

The various *Mycobacterium* species exhibit a great variation of growth rates, with doubling times ranging from a few hours to days. *M. tuberculosis* has a doubling time of between 15 to 24 hours, which is slow, when compared to one of the most regularly used laboratory bacteria for example, *Escherichia coli* that doubles every 20 to 30 minutes[31,32].

The MTBC is comprised of *Mycobacterium* species that are phylogenetically closely related to *M. tuberculosis* and are responsible for primary human infections or can be transmitted from infected animals. The MTBC comprises the aforementioned species, in addition to *M. caprae* and *M. pinipedii,* which are all closely genetically related, a fact highlighted by the efficacy of the BCG vaccine in protecting against human infections by the MTBC[33]. The *Mycobacterium* species that make up the MTBC, typically cause pulmonary TB in humans and exhibit a similar pathogenesis to *M. tuberculosis*[33]. The species of the MTBC are thought to be derived from a pool of ancestral tubercle bacilli, collectively called *Mycobacterium prototuberculosis*[34]. The MTBC species can be composed of two major lineages and emerged approximately 40,000 years ago, which is close to the proposed time of the migration of modern human populations from Africa[35,36]. Evidence of *M. tuberculosis* infection of humans from late Neolithic populations in Hungary, has recently been presented, which surpasses the previous oldest recorded TB cases identified in the mummified remains of Egyptians[8,37]. TB or illnesses that match the modern definition of the disease were recorded 3300 and 2300 years ago, in India and China, before the first exact pathological and anatomical description were presented by Francis Sylvius, in 1679 in his work *Opera Medica*[38]. The infectious origin of TB was first described by Benjamin Marten, in his publication “A new theory of Consumption”, in 1720[39]. The terms of consumption and phthisis were used to describe the disease throughout the 17th and 18th centuries, until the name tuberculosis was utilised by Johann Lukas Schonlein in the mid-19th century[40].

The causative agent of TB, *M. tuberculosis* (named in 1883), was first identified by the German doctor Robert Koch in 1882, utilising methylene blue staining, a technique that was further refined by both Koch and Paul Ehrlich, before finalisation of the approach by Ziehl and Neelsen in 1885[40]. Koch subsequently isolated the

tuberculosis bacillus an achievement that he received the Novel Prize for Medicine in 1905[41].

1.1.5 The *Mycobacterium* cell envelope

As previously described the *Mycobacterium* cell envelope is a defining feature of the bacteria, due to its highly-ordered, relatively impermeable nature, which is likely one of the key attributes that has contributed to both the longevity and persistence of the pathogen over time. Inhibition of cell wall biosynthesis and assembly have proven to be successful approaches for TB chemotherapy, with several drugs identified that are efficacious against *M. tuberculosis* growth. Subsequently, understanding the intricate *Mycobacterium* cell envelope and its biosynthesis has been a major research objective over the last decade.

The *Mycobacterium* cell envelope is comprised of core cell wall core, which consists of three primary components; an inner layer of mesh-like arrangement of peptidoglycan, which is covalently attached, through muramic acid, to middle section of highly branched arabinogalactan heteropolysaccharide that is in turn esterified, at its non-reducing ends to long-chain mycolic acids (Figure 1)[28].

Peptidoglycan, a complex glycopolymer, is a common component of most bacteria cell walls, sharing the same basic core structure of a glycan backbone and short cross-linked peptide side chains. The Peptidoglycan layer is the innermost section of the typical bacterial cell wall, interfacing with the inner cell membrane and has been shown to provide shape, rigidity and osmotic stability. Peptidoglycan is typically composed of repeating units of *N*-acetylglucosamine and *N*-acetylmuramic acid, which are cross-linked by short peptides. *Mycobacterium* peptidoglycan poses one feature that is unique; the oxidation of *N*-acetylmuramic acid to *N*-glycolylmuramic acid, which provides additional sites for hydrogen bonding and is believed to strengthen the structure. In addition, *Mycobacterium* peptidoglycan has been shown to act as a scaffold to support the rest of the complex cell envelope[42]. The peptidoglycan of bacteria can be broadly classified and mycobacterial glycopolymer belongs to the A1γ chemotype, as does that of *Escherichia coli, Agrobacterium tumefaciens, Salmonella typhi* amongst others[43]. *M. tuberculosis* tetrapeptides, which are comprised of L-Ala-γ-D-Glu-*meso*-A_2pm (or L-Lys)-D-Ala-D-Ala (A_2pm, 2,6-diaminopimelic acid) and become cross-linked to provide a mesh-like structure.

Another unique feature of Mycobacterial peptidoglycan is that it provides a site, the 6-position of some muramyl units, for attachment of the arabinogalactan layer[44].

The arabinogalactan layer of the cell wall is comprised predominantly of galactose (Gal) and arabinose (Ara) sugar residues, both in the furanose ring form, which are arranged in a highly branched macromolecule. The galactan chain is composed of alternating 5- and 6-linked D-galactanfurnosyl residues. In the *Mycobacterium* arabinogalactan, two thirds of the non-reducing termini pentaarabinosyl motifs are esterified with mycolic residues[44].

The outer mycolic acid layer of the *Mycobacterium* cell wall is comprised of long-chain α-alkyl-β-hydroxy fatty acids (C_{70-90}), which contribute both fluidity and permeability to the structure[27].

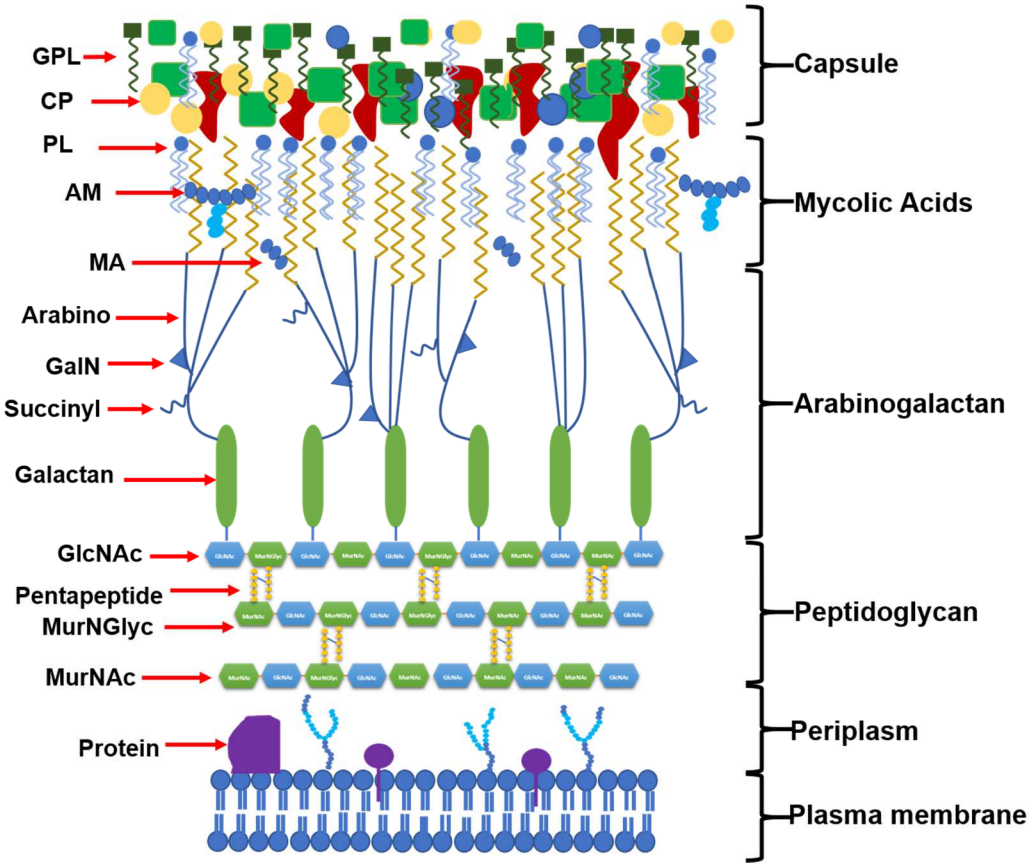


Figure 1. The *Mycobacterium* cell envelope.

A representative schematic of the *Mycobacterium* cell envelope, demonstrating the overall arrangement of the complex and highlighting most of the key components. GPL – glycopeptidolipids, CP – capsular polysaccharides, PL – phospholipids, AM – Arabinomannan, MA – Mannan. Image inspired by Keiser and Rubin[45], in addition to Chiaradia, *et al.*[46].

The cell wall is covered in two additional layers; an outer membrane, comprised of non-covalently linked lipids, lipoglycans and inert waxes, and a loosely attached capsule, made up of polysaccharides, proteins and a minor amount of lipids[29].

The lipids and lipoglycans are intercalated within the mycolic acid layer of the cell wall and are comprised of glycophospholipids; phosphatidyl-myo-inositol mannosides (PIMs) and their more related glycosylated end products, lipomannan (LM) and lipoarabinomannan (LAM). The glycophospholipids in tandem with the covalently linked mycolic acids form an outer membrane or "mycomembrane" that is a highly impermeable asymmetrical bilayer, which provides an additional barrier to small, polar molecules, such as external nutrients but also chemotherapeutic agents.

The final component of the *Mycobacterial* cell wall is a capsule of polysaccharide and protein, as well as a minor amount of glycolipid. The capsule has been postulated to play a key role in surviving and subsequently modulating the host immune response during infection[29].

1.1.6 Tuberculosis treatment and current drug therapies

TB is a treatable disease and has been for the last 46 years since the introduction of an inexpensive and effective quadruple drug therapy regimen. However, the current first-line TB treatment is challenging, primarily because accurate and early diagnosis is required. In addition, prior to prescription of the drug regimen, drug-resistance screening is required to prevent the spread of MDR- or XDR-TB. Following positive diagnosis, administration of effective treatment regimens takes at least 6 months, utilising "directly observed therapy" (DOT), followed by post-treatment monitoring and support. Directly observed therapy is an approach to administration of the anti-TB quadruple drug combination, which involves carers directly observing patients taking the treatment[47]. WHO estimates for the years from 2000 to 2017, are that TB drug treatment saved 54 million lives globally[14].

The first effective anti-TB drug discovered was streptomycin, in 1944, but during clinical studies antibiotic resistance was rapidly identified. This resistance issue was solved using the first combination therapy, which involved the inclusion of isoniazid and *para*-aminosalicyclic acid, but other problems were encountered, chiefly, patient adherence to the drug regimen. The discovery of rifampicin in 1963 was a major advance, as it enabled the development of an orally administered regimen that

reduced the length of therapy to 6 months. Rifampicin is a bactericidal antibiotic, as are isoniazid and pyrazinamide, which means that these drugs rapidly kill *M. tuberculosis*[48] (Figure 4).

The current recommended, first-line, TB treatment for drug-susceptible TB (DS-TB) is a 6-month regimen, which involves an intensive phase of a combination of rifampicin, isoniazid, ethambutol and pyrazinamide for 2 months, followed by a secondary, continuation phase of 4 months with the patient taking isoniazid plus rifampicin (also known as short-course chemotherapy). This drug regimen is implemented for all cases of DS-TB, including pulmonary, most cases of extrapulmonary and regardless of the patient's HIV status[49]

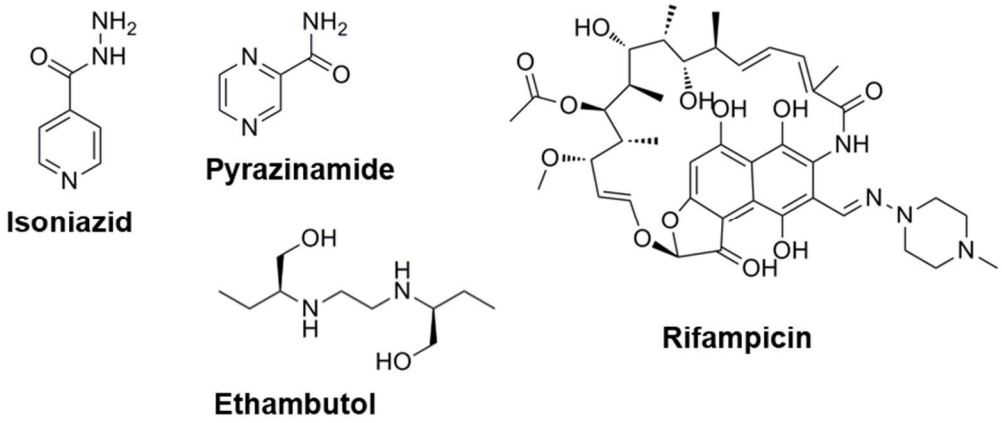


Figure 2. First-line TB drugs.

The WHO recommended first-line TB drugs[47].

The first-line TB antibiotics (Figure 2) all target different cellular pathways that include; disturbance of cell wall formation through inhibition of enoyl-[acyl-carrier-protein] reductase and arabinosyl transferases (isoniazid and ethambutol), disruption of transcription by inhibition of RNA polymerase (rifampicin) and acidification of the cytoplasm through binding RpsA, the ribosomal protein S1 (pyrazinamide)[18] (Table 1).

Table 1. First-line TB drugs and their effect on *M. tuberculosis*[a].

Drug	Abbreviation	Effect	Year of discovery (approval)[b]
Rifampicin, Isoniazid, Ethambutol, Pyrazinamide.	R or Rif, H or Inh, E or Emb, Z or Pza.	Inhibits transcription, Inhibits mycolic acid synthesis, Inhibits arabinogalactan biosynthesis Inhibits translation and *trans*-translation, acidifies cytoplasm.	1963 (1971), 1952 (1953), 1961 (1968), 1954 (1959).

[a] References for the information included in Table 1; Zumla*, et al.*[47], Barry[50]. [b] Year of approval quoted is the first reported global approval by any drug approval organisation.

There are several issues that can complicate the treatment of TB, especially when a patient is also HIV positive. The concurrent use of anti-retroviral therapy (ART) alongside first-line treatment of DS-TB, improves survival rates in HIV-infected individuals. However, treatment of TB in this patient population is complicated by potential drug-drug interactions and the risk of developing immune reconstitution inflammatory syndrome (IRIS). IRIS is a condition where the recovery of a HIV-positive patients immune system begins but then responds to an existing opportunistic infection, such as TB, with an overwhelming inflammatory response that conversely makes the symptoms of the infection worse[51]. Drug-drug interactions (DDI), which is defined as the co-administration of two drugs that subsequently have either a synergistic or antagonistic effect on one another. The new effect produced by the DDI is not present when either of the drugs are administered alone. The prevalent DDI when co-administering the first-line TB regimen and ART is between Rifampicin and HIV protease inhibitors. Rifampicin is a potent inducer of the hepatic cytochrome P450 3A (CYP3A) system, which are a family of monooxygenases that catalyse many reactions including drug metabolism, as well as synthesis of cholesterol, steroids and other lipids[52]. Induction of CYP3A by rifampicin leads to reduction of serum concentrations of HIV protease inhibitors, due to drug metabolism by these liver monooxygenases. Rifabutin, a related rifamycin drug to rifampicin, is the least potent inducer of CYP3A and is used as an alternative in the treatment of active TB in HIV-positive patients[53].

As previously discussed, MDR-TB is a growing global problem, and treatment of confirmed cases of these infections is complex. MDR-TB drug treatment would ideally involve bespoke regimens based on *in vitro* drug-susceptibility testing (DST) results for each patient's isolate[54]. The facilities required for DST, in addition to the lead time required to generate results, which takes from several weeks to several months, means that this approach is rarely tenable in real-world situations, especially in developing world countries. Subsequently, the choice of drugs to treat MDR-TB is determined by the pattern of drug resistance in the local geographical region, in combination with consideration of the patients TB drug history and underlying medical conditions. New medical diagnostics, such as Xpert MTB/RIF assay, that are both straight-forward to use and relatively cost-effective, are being made available to detect MDR-TB.

Treatment of both MDR/RR-TB is associated with a high rate of failure and relapse, partially due to the protracted treatment period, which can be from 20 to 28 months and must be administered daily under DOT. The treatment regimen, recently updated by the WHO, requires the use of a combination of several second-line drugs, which are divided in three different groups (Figure 3 and Table 2) [55].

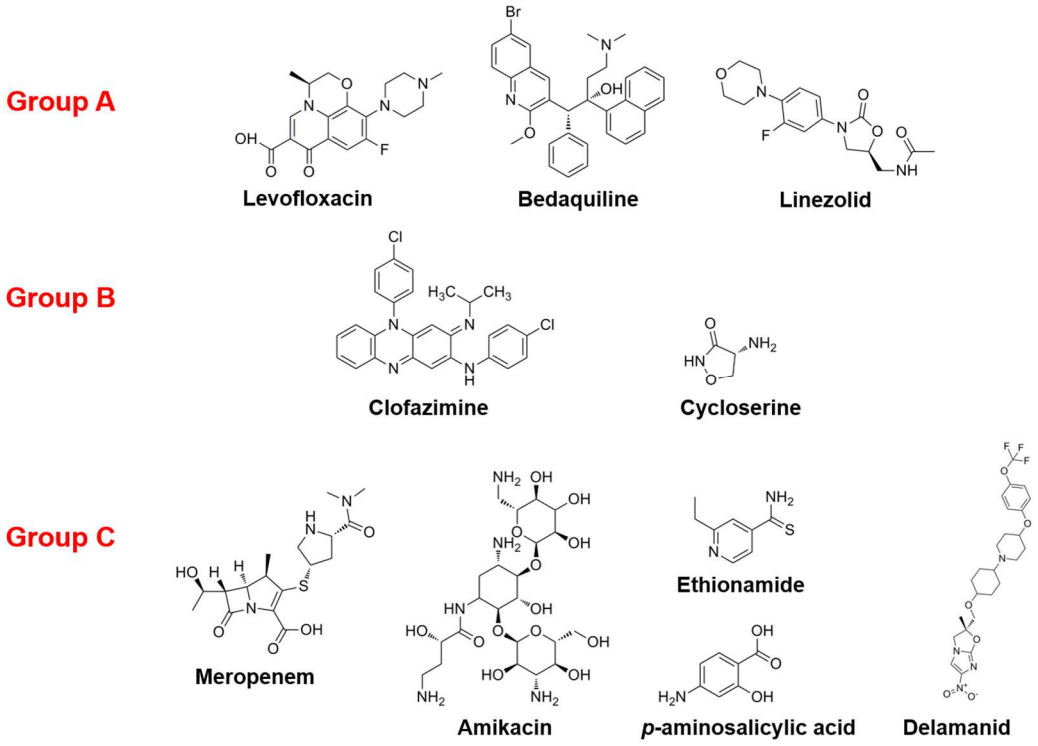


Figure 3. Second-line TB drugs.

Table 2. Medicines recommended for use in longer MDR-TB regimens.

Group	Medicines	Abbreviation	Year of approval [a]
Group A	Levofloxacin or Moxifloxacin	Lfx or Mfx,	1987 or 1988,
	Bedaquiline,	Bdq,	2012,
	Linezolid.	Lzd.	2000.
Group B	Clofazimine,	Cfz,	1986,
	Cycloserine or Terizidone.	Cs or Trd.	1968 or 1978.
Group C	Ethambutol,	E,	1968,
	Delamanid,	Dlm,	2014,
	Pyrazinamide,	Z,	1959,
	Imipenem-cilastatin OR Meropenem,	Ipm-Cln or Mpm,	1985 or 1996,
	Amikacin (or Streptomycin),	Am or (S),	1981 (or 1947),
	Ethionamide or Prothionamide,	Eto or Pto,	1968 or 1970,
	p-aminosalicylic acid	PAS.	1950.

[a] Year of approval quoted is the first reported global approval by any drug approval organisation.

The second-line medicines against MDR-TB include several drugs that inhibit protein synthesis by binding the 30S ribosomal subunit, including the aminoglycosides (streptomycin, amikacin) and Linezolid (an Oxazolidinone), that blocks the initiation of protein production. Other second-line drugs include D-cycloserine, which is a cell wall inhibitor, acting through D-alanine racemase and D-alanine ligase, as well as the fluoroquinolones (FQs; Moxifloxacin, Levofloxacin) that target *M. tuberculosis* DNA gyrase.[12] A number of medicines are no longer recommended, including the aminoglycosides capreomycin and kanamycin[55]. The removal of these medicines was due to an increased risk of treatment failure and relapse associated with their use in longer MDR-TB regimens[56].

The relatively, newly approved drugs bedaquiline and delamanid, are now recommended for routine use in treating MDR-TB, despite significant drug safety markers[57] (Figure 4).

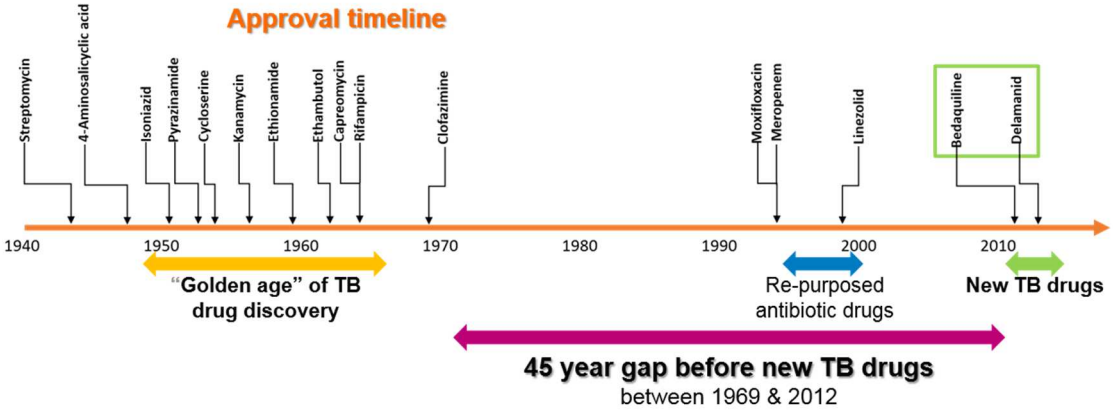


Figure 4. Approval timeline of antibiotics to treat *Mycobacterium tuberculosis.*

In order to try and solve compliance issues with the longer course MDR-TB treatment, a shorter MDR-TB treatment of 9 to 12 months is now recommended for all patients with pulmonary MDR and RR-TB that does not show resistance to second-line drugs[58]. Selection of the shorter MDR-TB treatment, is based on the feasibility of fully oral treatment regimens and the exclusion of a patients potential drug resistance[55].

The global success rates for treatment of patients with XDR-TB is 25%, falling to <20% when the drug resistance profile of the infection is beyond XDR. There is no standardised drug treatment regimen for XDR-TB, primarily due to the lack of effective and rapid diagnostics to identify afflicted patients. There are aspirational proposals for standardised drug regimens to treat both pre-XDR-TB and XDR-TB to rectify the current low success rates. However, the requirements for the proposed regimen are the inclusion of at least four new antibiotic drugs and at least two of these inhibitors would need to be "core" drugs. These requirements would require the use of new drugs, such as bedaquiline and delamanid, as core drugs, which given the propensity of these treatments to cause adverse events, is far from ideal[59].

There have been several reports of drug-resistant *M. tuberculosis* strains that have resistance beyond the accepted classification for XDR-TB, which have been identified in countries in such as Italy, Iran, India and China[60]. These incidences of drug-resistant TB have been termed either extensively drug-resistant TB (XXDR-TB) or totally drug-resistant TB (TDR-TB) and are broadly defined by resistance to additional or all second-line drugs. The DST methods for identifying resistance beyond XDR-TB are problematic, in addition to the paucity of data for the reported

cases, are the primary reasons for the WHO not officially recognising and standardising the classification of these drug resistant cases of TB[61].

The cost of treating a TB infection is not insignificant, with the treatment of DS-TB in 2014 approximated to be US$14,659 (£11,353) in high-income countries and US$258 (£200) in low-income countries. The cost of MDR-TB treatment in 2014 was US$83,365 (£64,567) in high-income countries and US$1,281 (£992) in low-income countries[62]. The global median cost per patient treated for DS-TB in 2017 was US$1,224 (£935) and the range is between US$85 to US$20,00 (£65 to £15,277). The global cost of the drug regimen to treat MDR-TB is US$7,141 (£775) per person and the range is between US$750 to US$60,00 (£572 to £45,833)[14]

In 2015, a study by the "Post-2015 Consensus" reported for every dollar invested in the treatment of TB, there is a return of US$43 (£33) and investment in other countries programmes can also lead to beneficial domestic returns[63]. Despite these reports and the associated evidence, since 2009, there has been a significant under spend, against the targets set out by the WHO, on investment in global TB control[64]. In the 2018 WHO TB report, a funding gap of US$3.5 billion (£2.67 billion) for TB prevention and care, a deficit that will increase to US$5.4 billion (£4.12 billion) by 2020 if current funding levels are maintained, while a gap of US$1.3 billion (£ 0.99 billion) for TB research and development[14].

The current TB treatment success rate for DS-TB is 83%, however, for MDR-TB, the success rate was only 55% and for XDR-TB only 25%. There has been significant progress been made over recent years with approval of the first new TB drugs, bedaquiline and delamanid, in more than 40 years[65,66]. In addition, there are currently five new monotherapy drugs in the late phases of clinical development for the treatment of TB[67]. There are also several new vaccines and diagnostic techniques under development[68,69]. Despite the progress being made for TB treatment, prevention and care, there are still may challenges, including the poor understanding of LTBI, as well as the year-on-year rise in the number of MDR and XDR-TB cases and finally, the safety of some of the current treatments. Consequently, developing new drugs that are effective against both latent and active-TB, and that can shorten and simplify the current regimens and treatment periods constitute the major objective in modern anti-tuberculosis drug discovery.

1.2 Acetyltransferases and the Left-handed beta helix sub-family

1.2.1 Acetyltransferases, an introduction

Acetyltransferases (or transacetylases) are enzymes that transfer an acetyl group, to form either esters or amides and are more broadly classified as acyltransferases (EC 2.3)[70]. Acetyltransferases, which release CoA-SH as one of the products of catalysis, perform many diverse functions in nature, from post-translation modification of histone *N*-terminal tail domains in vertebrates and yeast[71], to drug resistance in bacteria[72]. There are a wide range of acetyltransferase enzyme structures and perhaps the best characterised are the GCN5-related *N*-acetyltransferase (GNAT) superfamily[73]. Members of this superfamily have a structurally conserved protein fold, which is comprised of an N-terminal strand followed by two helices, three antiparallel *β* strands, followed by a central helix, a fifth *β* strand, a fourth α helix, and a final *β* strand. Well characterised exemplars of the GNAT superfamily are the histone acetyltransferases (HAT), including GCN5, PCAF and CBP, in addition to the bacterial resistance factors, the aminoglycoside acetyltransferases. Distinct acetyltransferase enzymes families, include amongst others the MYST (MOZ, Ybf2/Sas3, Sas2, Tip60) family, which catalyses histone acetyltransferase[74], and the arylamine *N*-acetyltransferases (NATs)[75]. The *N*-acetyltransferase enzymes, along with varying structural folds, catalyse acetyl transfer by several diverse mechanisms including bi-bi ping-pong and ternary complex, which can either be ordered or random[73,76].

N-acetyltransferases are defined as enzymes that catalyse the transfer of an acetyl group from Ac-CoA to a primary amine of a substrate. In addition, there are enzymes that catalyse *O*-acetyl and *S*-acetyl transfer, which are defined as transferases that transfer an acetyl group from Ac-CoA to either a hydroxyl, or a thiol group of the given substrate[77].

1.2.2 Left-handed beta-helix family of acetyltransferases

The left-handed *β*-helix (L*β*H) superfamily is a distinct class of acyl- and acetyltransferases, which includes the following enzymes; serine acetyltransferase (SAT)[78], galactoside acetyltransferase (GAT)[79], tetrahydrodipicolinate *N*-

succinyltransferase (DapD)[80], maltose acetyltransferase (MAT)[81], xenobiotic acetyltransferase[82], UDP-GlcNAc acyltransferase (LpxA)[83] and UDP-3-*O*-acyl-glucosamine *N*-acyltransferase (LpxD)[84]. The members of this family of enzymes are involved in a variety of cellular processes, including amino acid metabolism, cell wall biosynthesis and detoxification of foreign agents, such as antibiotics and herbicides. The L*β*H enzymes are found predominantly in higher plants and microorganisms and also include non-transacetylases, such as γ-class carbonic anhydrase from the methanogenic archaeon *Methanosarcina thermophila* (pdb 3OTZ)[85].

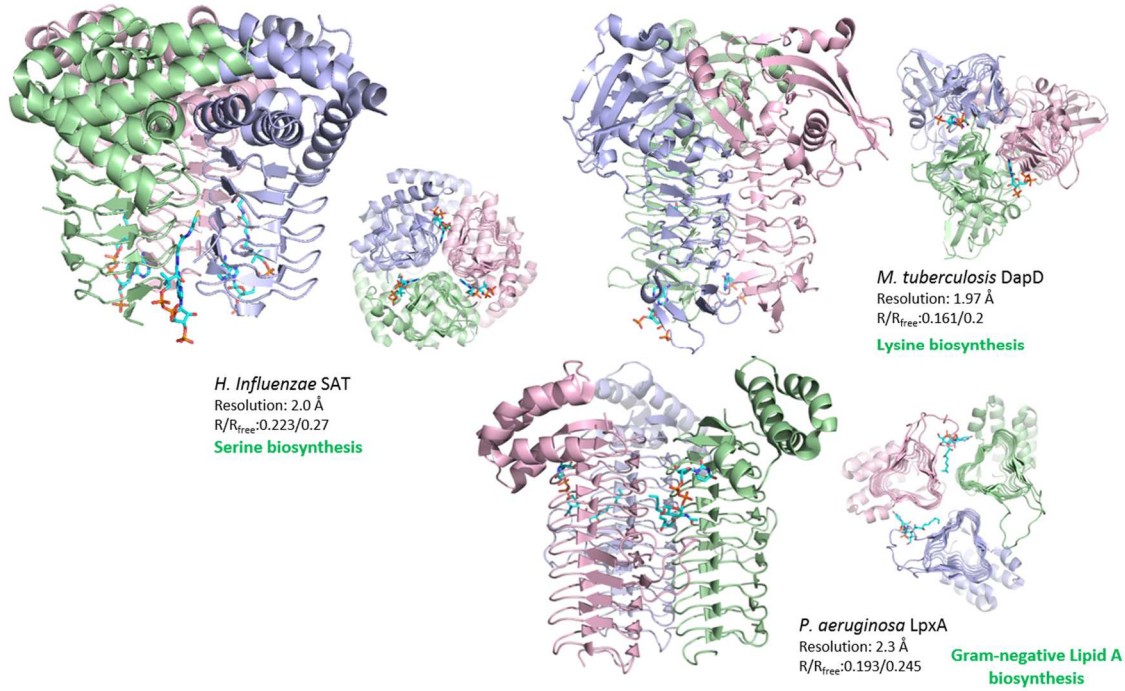


Figure 5. Example structures of L*β*H acyl- and acetyltransferases.

The structures of the LH acyl- and acetyltransferases shown in this figure can be found in pdb, using the following references: SAT, 1SST; DapD, 3FSY and LpxA, 5DG3.

The L*β*H acyl- and acetyltransferase family is characterised by a left-handed *β*-helix tertiary protein fold, a structure that is an equilateral triangle in cross-section (Figure 5)[86]. The tertiary structure is formed by parallel *β*-strands folding into a helix with three strands, which make up each of the sides of the triangle, per turn. The sequence of the *β*-helix is often recognised by a hexapeptide repeating sequence, comprising two residues in a *β*-strand, followed by four residues in a left handed 120° turn. These turns are tight and flat, which act to orient the central peptide plane

perpendicular to the plane of the turn. The core of the fold predominantly comprises of bulky, hydrophobic residues (Ile, Val) and generally, corner stacks of smaller residues (Ala, Thr). The consensus sequence for this imperfect, hexapeptide repeat is [Leu/Ile/Val]-[Gly/Ala/Glu/Asp]-X-X-[Ser/Thr/Ala/Val]-X, which can also be referred to as an Ile patch[87,88]. The narrow central channel that runs collinear to the axis of the prismatic L*β*H domain is too narrow to accommodate solvent[89]. The known L*β*H acyl- and acetyltransferase enzymes that have been structurally characterised using X-ray crystallography are listed in Table 3.

Table 3. Structurally characterised L*β*H acetyl- and acyltransferases[a].

L*β*H enzyme	Abbrev. (ref.)	Species	Acyl-thioester[b]	Acceptor substrate	pdb code
Serine acetyltransferase	SAT[90]	*E. coli*	Ac-CoA	Ser	1T3D
Serine acetyltransferase	SAT[90]	*H. Influenzae*	Ac-CoA	Ser	1SSM
Serine acetyltransferase	SAT[91]	*Glycine max*	Ac-CoA	Ser	4N6A
Galactoside acetyltransferase	GAT[79]	*E. coli*	Ac-CoA	galactopyranosides	1KQA
Galactoside acetyltransferase	GAT[92]	*Staphylococcus aureus*	Ac-CoA	Maltose, glucose	3V4E
Maltose acetyltransferase	MAT[81]	*E. coli*	Ac-CoA	Maltose, glucose	1OCX
Virginiamycin acetyltransferase D	VatD[93]	*Enterococcus faecium*	Ac-CoA	Virginiamycin	1KHR
Xenobiotic acetyltransferase	PaXAT[82]	*Pseudomomas aeruginosa*	Ac-CoA	Chloramphenicol	1XAT
UDP-GlcNAc acyltransferase	LpxA[86]	*E. coli*	ACP	UDP-GlcNAc	1LXA
UDP-GlcNAc acyltransferase	LpxA[94]	*Heliobacter pylori*	ACP	UDP-GlcNAc	1J2Z

UDP-GlcNAc acyltransferase	LpxA[95]	*Leptospira interrogans*	ACP	UDP-GlcNAc	3HSQ
UDP-GlcNAc acyltransferase	LpxA[96]	*Acinetobacter baumannii*	ACP	UDP-GlcNAc	4E6T
UDP-GlcNAc acyltransferase	LpxA[97]	*Arabidopsis thaliana*	ACP	UDP-GlcNAc	3T57
UDP-GlcNAc acyltransferase	LpxA[98]	*Burkholderia thailandensis*	ACP	UDP-GlcNAc	4EQY
UDP-GlcNAc acyltransferase	LpxA[99]	*Francisella novicida*	ACP	UDP-GlcNAc	5F42
UDP-GlcNAc acyltransferase	LpxA[100]	*P. aeruginosa*	ACP	UDP-GlcNAc	5DEM
UDP-GlcNAc acyltransferase	LpxA[101]	*Bacteroides fragilis*	ACP	UDP-GlcNAc	4R36
UDP-3-*O*-acyl-glucosamine *N*-acyltransferase	LpxD[102]	*E. coli*	ACP	UDP-3-*O*-acyl-glucosamine	3EH0
UDP-3-*O*-acyl-glucosamine *N*-acyltransferase	LpxD[84]	*Chlamydia trachomatis*	ACP	UDP-3-*O*-acyl-glucosamine	2IU8
UDP-3-*O*-acyl-glucosamine *N*-acyltransferase	LpxD[96]	*A. baumannii*	ACP	UDP-3-*O*-acyl-glucosamine	4E79
UDP-3-*O*-acyl-glucosamine *N*-acyltransferase	LpxD[103]	*P. aeruginosa*	ACP	UDP-3-*O*-acyl-glucosamine	3PMO
Tetrahydrodipicolinate (THDP) succinyltransferase	DapD[89]	*E. coli*	Succinyl-CoA	THDP	1TDT
THDP succinyltransferase	DapD[104]	*M. tuberculosis*	Succinyl-CoA	THDP	3FSX
THDP succinyltransferase	DapD[105]	*P. aeruginosa*	Succinyl-CoA	THDP	3R5D

THDP succinyltransferase	DapD[106]	*Corynebacterium glutamicum*	Succinyl-CoA	THDP	5E3P
3-acetamido-3,6-dideoxy-α-D-glucose acetyltransferase	QdtC[107]	*Thermoanaerobacterium thermosaccharolyticum*	Ac-CoA	3-Acetamido-3,6-dideoxy-R-D-glucose	3FS8
UDP-GlcNAc3NA acetyltransferase	WlbB[108]	*Bordetella pertussis*	Ac-CoA	UDP-GlcNAc3NA	3MQH
3-acetamido-3,6-dideoxy-α-D-galactose acetyltransferase	FdtD[109]	*Shewanella denitificans*	Ac-CoA	dTDP-Fuc3N	4MZU
UDP-4-amino sugar acetyltransferase	PglD[110]	*Campylobacter jejuni*	Ac-CoA	UDP-4-amino sugar[c]	3BSW
UDP-4-amino sugar acetyltransferase	PglB[111]	*Neisseria gonorrhoeae*	Ac-CoA	UDP-4-amino sugar[c]	4M98
UDP-4-amino sugar acetyltransferase	WeeI[111]	*A. baumannii*	Ac-CoA	UDP-4-amino sugar[c]	4M9C
Virginiamycin acetyltransferase A	VatA[112]	*S. aureus*	Ac-CoA	Virginiamycin	4MYO
Perosamine *N*-acetyltransferase	PerB[113]	*Caulobacter crescentus*	Ac-CoA	GDP-N-acetylperosamine	4EA7
GlcNAc-1P uridyltransferase	ST0452[114]	*Sulfolobus tokodaii*	Ac-CoA	GlcN-1P	5Z09
dTDP-4-amino-4,6-dideoxyglucose *N*-acetyltransferase	AntD[115]	*Bacillus cereus*	β-hydroxybutyryl-CoA	dTDP-sugar[d]	3VBM

[a] LβH enzymes only listed when an accompanying manuscript is published. [b] R-3-hydroxyacyl acyl carrier protein (ACP). [C] UDP-2acetamido-4-amino-2,4,6-trideoxyglucose (UDP-4-amino sugar). [d] dTDP-4-amino-4,6-dideoxyglucose (dTDP-sugar).

The crystal structures of the LβH containing enzymes show that all family members are arranged as trimers composed of three identical subunits that are related to one another by a crystallographic three-fold rotation axis. Each subunit is typically

composed of two domains; an amino-terminal domain of variable length consisting a mixture of *α*-helices, hairpin loops and *β*-sheet linked to a carboxy terminal domain containing predominantly coiled parallel *β*-sheet structure arranged as a L*β*H. The length of the L*β*H domain varies for each of the family members; the shortest reported is xenobiotic acetyltransferase from *Pseudomonas aeruginosa* (*P. aeruginosa,* PaXAT), which has five coils, compared to SAT (six coils), LpxA (ten coils), DapD (seven coils), and both *E. coli* and *Chlamydia trachomatis* LpxD (ten coils) (Figure 6). The L*β*H containing proteins form an equilateral-triangular prism quaternary structure, with the subunit-subunit interface, comprising 65% nonpolar residues and extends for almost the full length of the trimer (Figure 6 (A), (B)). The trimer interface occludes a significant proportion (up to one third) of each subunit's accessible surface area[84]. Each subunit of the L*β*H containing enzymes typically has an extended loop, which does not obey the hexapeptide repeat sequence pattern (Figure 6 (C)). These extended loops are thought to represent a means of endowing structural and functional diversity on the hexapeptide family members, as the coiled L*β*H domain is largely invariant. Indeed, the three active sites of the L*β*H homotrimers that are created near the inter-subunit regions are completed by contributions from the neighbouring subunits extended loop. The extended loops form a tunnel-like structure over the active site and contribute residues that are involved in both substrate recognition and catalysis[116].

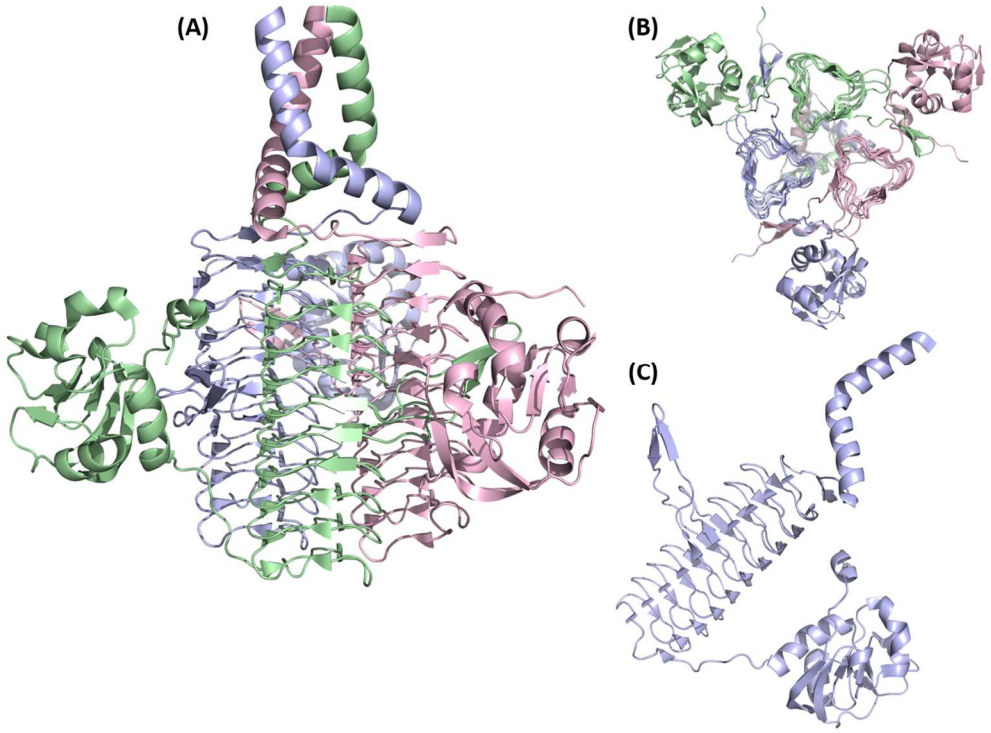


Figure 6. Structure of *P. aeruginosa* LpxD, a L*β*H acyltransferase.

This figure depicts the *P. aeruginosa* LpxD (pdb 3PMO) quaternary structure. Each monomer is displayed as a cartoon representation of the secondary structure. (A) Side view of the trimer, (B) bottom view of the LpxD trimer and (C) representation of the LpxD monomer, which clearly shows the ten coils of the L*β*H domain.

The site of Ac-CoA binding to the L*β*H acetyltransferase trimers is similar and involves Ac-CoA adopting a bent confirmation, with the glycosidic linkage orientated parallel to the extended pantothenic acid region. Typically, Ac-CoA adopts the same confirmation to all three binding sites of the acetyltransferase trimer. The carbonyl oxygen of the acetyl group of Ac-CoA is presented to the second substrate binding site and typically hydrogen bonded to two active site residues, invariably one of these amino acids is a His^{112}. Small hydrophobic or polar residues such as Ala, Val, Ser or Thr are packed around the pantothenic acid of Ac-CoA and form numerous interactions. The diphosphate negative charges of the pantothenic acid are balanced by interactions with basic residues and the adenine moiety of CoA-SH is stabilised by interactions with a conserved Pro^{81}.

The second substrate binding site is found at the alternative end of the active site tunnel to the CoA-SH binding site. Second substrate specificity of the L*β*H acetyltransferases is conferred by the surrounding residues. GAT and MAT have different hexose substrate specificity; GAT has an Asp residue that makes an

interaction to the hexose ring C3 and C4 hydroxyls of galactosyl moieties, whereas MAT has a Val residue in the corresponding position[81]. The catalytic centre of the active site is found where the tunnel formed by the L*β*H extended loop meets the second substrate binding site. The active site, catalytic centre of the L*β*H containing enzymes generally have at least one histidine residue present.

The non-acetyl- and acyltransferase enzymes that share the L*β*H structural motif, arranged as a homotrimer are the aforementioned γ-class carbonic anhydrases, which are zinc-containing enzymes that catalyses the interconversion of CO_2 and HCO_3-. There are seven genetically distinct classes of carbonic anhydrases (Cam): *α, β*, *γ, δ, ζ, η* and *ɵ,* which share very little sequence and three-dimensional similarity[117,118]. The structure of *γ*-Cam polypeptide is a L*β*H, in cross section resembling an equilateral triangle, which has seven coils and has a short *α*-helix at the carboxy-terminal. The *γ*-Cam monomer forms a homotrimer, arranged around a 3-fold axis of symmetry, with approximately 25% of the molecular surface per subunit involved in trimerisation[119]. The active site is formed at the interface between two subunits, with residues from both polypeptides involved in substrate binding and catalysis. Trimerisation leads to formation of three active sites per homotrimer, each of which can bind a Zn^{2+} ion, coordinated by the three imidazole side chains of His residues. There have been several *γ*-Cam L*β*H enzymes reported from bacteria such as *Thermosynechococcus elongatus* (pdb 3KWC)[120] and *E. coli* (pdb 3TIO)[121]. Non-acetyl- and acyltransferases L*β*H enzymes, that are also not classified as *γ*-Cam homotrimers have also been reported, including RicA, a metalloprotein from *Brucella abortus* that binds human small GTPases (pdb 4N27)[122] and a putative Cam, Ferripyochelin binding protein from *Pyrococcus horikoshii* that in addition to binding three Zn^{2+} ions, also binds a Ca^{2+} at the trimer interface (pdb 1V67)[123].

1.2.3 Kinetic and chemical mechanisms of L*β*H acyltransferases

The consensus kinetic mechanism reported for the L*β*H acetyl- and acyltransferases is a sequential bi-bi ternary complex mechanism, which has been reported for orthologues of SAT, GAT, MAT and DapD. The mechanism is typically equilibrium ordered with Ac-CoA or acyl-CoA binding to the enzyme first, followed by the second substrate. Product release is also ordered with the acetylated or acylated substrate released first, followed by CoA-SH. The kinetic mechanisms have been determined

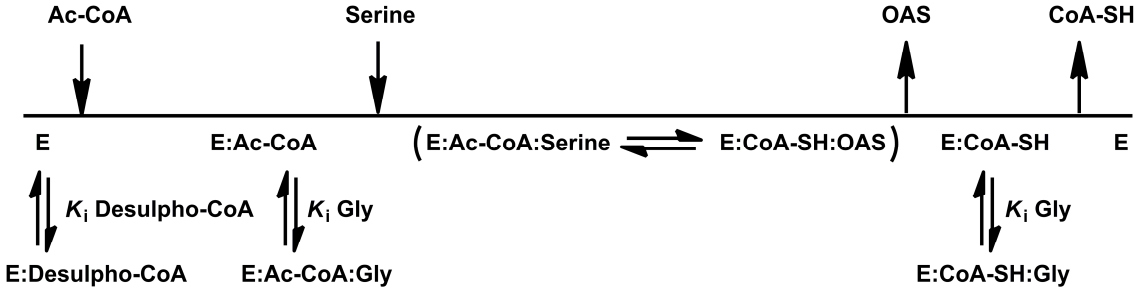
by use of steady state techniques, primarily absorbance spectroscopy, to carry out initial velocity pattern studies, as well as product and dead-end inhibition pattern experiments. Steady state constants and kinetic mechanism data for exemplars of L*β*H enzymes are presented in Table 4.

Table 4. Steady state parameters and kinetic mechanisms of L*β*H acetyl- and acyltransferases.

L*β*H enzyme	Species	k_{cat} (s^{-1})	$K_{m,\ Ac\text{-}CoA}$ (mM)	$K_{m,\ 2}$ [a] (mM)	kinetic mechanism
SAT [b]	*H. influenzae*	1350 ± 250	0.7± 0.1	23 ± 4	sequential ordered bi-bi
SAT [b]	*Salmonella typhimurium*	135 ± 5	0.4 ± 0.03	3.2 ± 0.2	sequential ordered bi-bi
GAT [c]	*E. coli*	14.7	0.104	63.4	sequential ordered bi-bi
MAT [d]	*E. coli*	38.5	0.0107	17.2	N.D.
DapD [e, k]	*E. coli*	N.D.	0.015	0.02	ternary complex formation, bi-bi
VatA [f]	*S. aureus*	8.84	N.D.	0.124 ± 0.018	sequential ordered bi-bi
PglD [g]	*C. jejuni*	7.05 ± 1.0 x 10^5	0.295 ± 0.002	0.274 ± 0.006	sequential ordered bi-bi
PglB [h]	*N. gonorrhoeae*	6.1 ± 0.7 x 10^4	0.286 ± 0.035	0.099 ± 0.007	sequential ordered bi-bi
WeeI [i]	*A. baumannii*	3.2 ± 0.06 x 10^5	0.078 ± 0.028	2.52 ± 0.54	sequential ordered bi-bi
PerB [j]	*C. crescentus*	304.5	0.15 ± 0.02	0.087 ± 0.011	sequential ordered bi-bi

[a] Acetyl- acceptor substrate. The following acetyl-acceptor substrates were used for the respective studies; [b] Ser, [c] isopropyl-β-D-thiogalactopyranoside (IPTG), [d] isopropyl β-D-thio-glucose, [e] tetrahydrodipicolinate, [f] virginiamycin M1, [g] UDP-4-amino sugar, [h] UDP-4-amino sugar, [i] UDP-4-amino sugar and [j] GDP-perosamine. [k] Succinyl-CoA was used as the alternative acyl-CoA.

The most comprehensive study of L*β*H acetyltransferase kinetic mechanism was carried out for *H. Influenzae* SAT by Johnson, *et al.* from the Cook research group, at University of Oklahoma[116]. Initial velocity patterns were carried out for SAT acetyl transfer to Ser by varying L-serine at several fixed concentrations of Ac-CoA in assay buffer adjusted to either pH 6.5 or 7.5. The initial velocity pattern at the pH 6.5 conditions, when plotted on a Lineweaver-Burke plot was best described by a series of intersecting lines converging on the ordinate. These data contrast with the initial velocity pattern data in the pH 7.5 conditions, varying both substrates and using the same analysis procedure, which was described as a series of intersecting lines that appear to be close to parallel. Thorough product and dead-end inhibition studies were carried out to ascertain the order of substrate binding and product release. Desulpho-CoA was utilised as a dead-end analogue of CoA-SH and Ac-CoA, and subsequent inhibitor pattern studies were competitive against both substrates in the direction of L-serine acetylation. Desulpho-CoA was competitive against CoA-SH and noncompetitive against *O*-acetyl-L-serine (OAS) in the direction of CoA-SH acetylation. Uncompetitive inhibitor patterns were identified when titrating the dead-end analogues Gly and *S*-methyl- L-serine (SMC) against Ac-CoA and CoA-SH, respectively. The *H. Influenzae* SAT product and dead-end inhibition studies were consistent with a sequential ordered bi-bi catalytic mechanism, represented by Scheme 1.



Scheme 1. Proposed kinetic mechanism for *H. Influenzae* SAT acetyltransferase activity.

Schematic depiction of the kinetic mechanism of SAT, adapted from Johnson, *et al.*[116]

The ordered nature of the kinetic mechanism for *H. Influenzae* SAT confirms earlier observations, in 1994, made by the same research group when studying *S.*

typhimurium SAT. These early studies, characterised that the *S. typhimurium* SAT catalysed acetyl transfer to L-serine by a bi-bi ping-pong kinetic mechanism, which is in contrast to the *H. Influenzae* enzyme[124].

The rate-limiting steps of acetyl transfer for L*β*H acetyl- and acyltransferases has been studied by solvent kinetic isotope effects (SKIEs). The *H. Influenzae* SAT SKIEs on *V* and V/K_{ser} are 1.9 and 2.5, respectively, in addition there is a linear dependence of both kinetic parameters on the $\%D_2O$, which is indicative of the involvement of a single proton transfer during catalysis being rate-determining. The reaction step that was the rate-determining step is catalysis after ternary complex formation, indicated by the effect on V/K_{ser}. Nucleophilic attack by the serine hydroxyl on the acetyl thioester is likely limiting, with a general base accepting a proton in this reaction. The difference between *V* and V/K_{ser} shows that the steps accompanying proton transfer are not solely rate limiting[125]. Rate limiting chemistry for the L*β*H enzymes has not been widely studied, particularly using SKIE, however, there are some reports for GAT that the primary rate-determining step is either the loss of the acetylated acceptor from the binary complex or the interconversion of the substrate and the product ternary complexes[126].

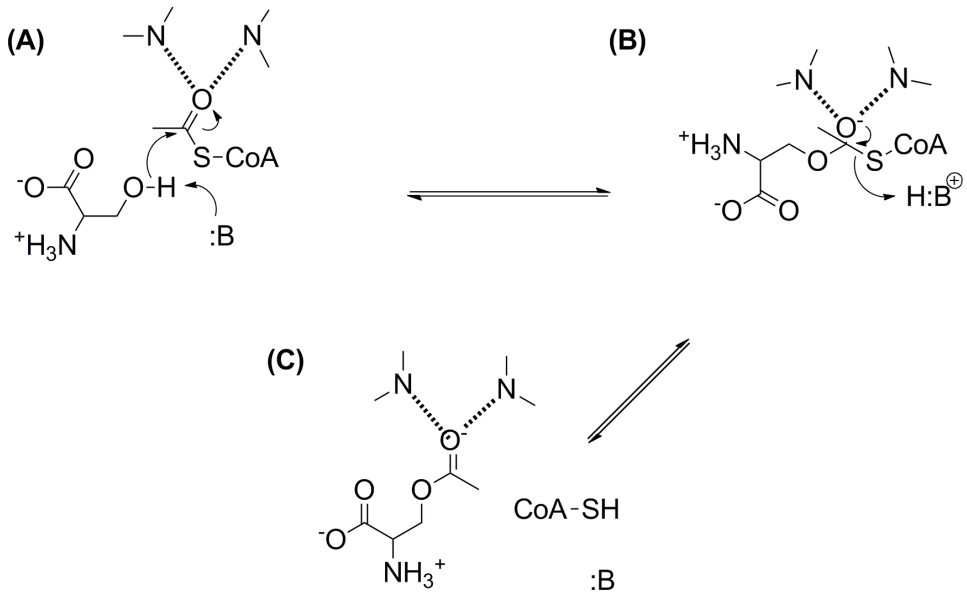
The catalytic mechanism of L*β*H enzyme catalysed acetyl or acyl transfer has been widely studied, in some cases utilising pH dependence studies to characterise the protonation state of the enzyme and reactant functional groups required for protein confirmation, binding and catalysis. Determination of the pH dependence of *V/K* reflects either the protonation state of titratable groups on the free enzyme or reactant. The protonation states of the titratable groups, identified in the *V/K* profile, can be important for either substrate binding, catalysis or both. Conversely, the pH dependence of *V* is indicative of the enzymic titratable groups that are required for catalysis. The pK_a values identified in both *V/K* and *V* profiles for the L*β*H acetyl- and acyltransferases are summarised in Table 5.

Table 5. Exemplar pH profile study data for L*β*H acetyltransferases.

LβH enzyme	fixed substrate	varied substrate	*V/K* pK_{a1}	*V/K* pK_{a2}	*V* pK_{a1}	pH range
SAT [a]	Ac-CoA	Ser	7.2 ± 0.2	N.D.	6.8 ± 0.2	5.5 – 10.0
SAT [a]	Ser	Ac-CoA	6	N.D.	6.8 ± 0.2	5.5 – 10.0
PerB [b]	Ac-CoA	GDP-perosamine	7.8 ± 0.3	9.3 ± 0.2	N.D.	6 - 11
PglD [c]	Ac-CoA	UDP-4-amino sugar	7.5 ± 0.1	9.5 ± 0.4	N.D.	6.5 – 9.0

[a] *H. Influenzae* SAT[127]. [b] *Caulobacter crescentus* PerB[113]. *Campylobacter jejuni* PglD[110].

A rigorous study of L*β*H acetyltransferase enzyme chemical mechanism was carried out for *H. Influenzae* SAT, utilising pH-rate studies on both initial velocity patterns and dead-end analogue inhibitor patterns, in addition to the aforementioned SKIE experiments[127]. A subset of the pH-rate study data for *H. Influenzae* SAT is presented in Table 5. These pH-rate studies lead to determination of a pK_a of ~7 for a single enzyme residue that must be unprotonated for optimum activity. In addition, the K_i values for the competitive dead-end inhibitors; Gly and L-Cys, are pH dependent. These data, in combination with SKIE results, led to the proposal of a general base catalytic mechanism, which is represented in Scheme 2.



Scheme 2. Proposed chemical mechanism for *H. Influenzae* SAT acetyltransferase activity.

Schematic depiction of the chemical mechanism of SAT, adapted from Johnson, et al.[127]. As previously discussed both substrates need to be bound prior to acetyl transfer. (A) Nucleophilic attack of the oxygen from the Ser side chain hydroxyl as it attacks the carbonyl carbon of Ac-CoA. General base accepts the proton from the side chain hydroxyl as it attacks the carbonyl of Ac-CoA. It is proposed that the carbonyl oxygen is stabilised by positive dipoles from backbone nitrogens (identified by hashed bonds, which all represent the nitrogen H atoms). The nucleophilic attack of the Ser oxygen on the Ac-CoA carbonyl results in a tetrahedral intermediate. (B) The same residue that served as a general base then acts as a general acid and donates a proton to the sulphur atom of CoA-SH. (C) The tetrahedral intermediate collapses after a thiolate is expelled, giving products OAS and CoA-SH.

The identity of the *H. Influenzae* SAT residues involved in catalysis was subsequently identified by SDM studies as His154, Asp139 and His189[128]. A catalytic dyad linkage was identified between His154 and Asp139 and His154 was proposed to be the likely general base, as mutation of His154 to Asn (H154N) significantly reduced the k_{cat}. Catalysis was not completely ablated by the H154N SDM and His189 was proposed as a backup general base, albeit at a much lower efficiency than His154. Asp139, through dyad linkage with His154, is proposed to facilitate catalysis by increasing the basicity of the sidechain imidazole of His154[128].

A similar chemical mechanism was proposed for PglD from *C. jejuni*, with His125, Glu126, Glu124, Asn118 and His134 put forward as the key catalytic residues. Glu126, which is one half of a catalytic dyad, serves to increase the basicity of the imidazole sidechain of the proposed general base His125. Asn118 and His134 are

believed to hydrogen-bond with the thioester carbonyl oxygen of Ac-CoA. The catalytic mechanism was proposed to proceed in the same manner as SAT, with a key difference, which is the general base does not subsequently act as a general acid. The proposed general acid in the PglD active site was Glu124 that donates a proton to His125 to regenerate the general base to its pre-catalytic state[110]. The L*β*H enzyme, PerB, is also proposed to utilise a general base mechanism to catalyse acetyl transfer, with His141 acting as the catalytic residue. The other key PerB catalytic residues were proposed to be Asp142, which was postulated to perform a similar role as the PglD residue Glu126 and Gly159 that was suggested to provide an oxyanion hole for stabilisation of the tetrahedral intermediate[113]. PerB is proposed to bind the substrate, GDP-perosamine, in the active site with a neutral amino group, which was also a mechanism that has been postulated for PglD[129].

There are members of the L*β*H family that are proposed to catalyse acetyl- or acyl transfer by an alternative chemical reaction mechanism, that does not use a catalytic histidine residue as a general base. Exemplars of these L*β*H acetyl- or acyltransferases are WlbB, QdtC and AntD[113], which all lack active site residues that are able to function as a general base. WlbB and QdtC have been shown to bind their sugar substrates in a nearly identical manner and in stark contrast to that observed for PglD[108]. The proposed chemical mechanism for WlbB, QdtC and AntD, in addition to a further sub-family member FdtD[109], was first proposed for QdtC by Thoden *et al.* in 2009[107]. Sequence alignment with PglD identified a conserved active site His, QdtC residue His123, as the likely candidate for the catalytic general base. However, this His residue was ~ 8 Å from the sugar substrate amino group and coupled with observed differences with sugar binding in PglD versus QdtC, led to concerns that there was an artefact in the QdtC crystal structures. SDM studies identified that His125 mutations to either Asn or Ala did not affect the catalytic efficiency of QdtC and confirmed that His125 did not act as a general base. A chemical mechanism for QdtC was proposed that requires the nucleotide-linked sugar to bind in the active site in a unprotonated form. The nucleotide-linked sugar is then proposed to hydrogen bond with the carboxamide group of Asn159, which aligns the amino nitrogen of the substrate in the correct orientation for nucleophilic attack on the carbonyl carbon of Ac-CoA. The other hydrogen on the nucleotide-linked sugar amino group is directed at the sulphur of Ac-CoA, while the lone pair of electrons on the amino nitrogen is directed at the carbonyl carbon of Ac-CoA. As the

amino nitrogen attacks, the bond between the carbonyl carbon and the sulphur of Ac-CoA becomes longer and eventually breaks. The sulphur of Ac-CoA is essentially proposed to serve as the catalytic base by accepting a proton from the sugar amino group. Subsequently, based on the difference between chemical mechanisms and also substrate binding orientations, it was proposed that two subfamilies of *N*-acetyltransferases that acetylate nucleotide-sugars existed in the L*β*H superfamily. Class 1 enzymes, such as PglD, PerB, as well as VatA, utilise a conserved His residue as a catalytic base to remove a proton for the amino group of the sugar substrate, prior to nucleophilic attack on the acetyl group of Ac-CoA. This sub-family of the L*β*H superfamily have a chemical mechanism that is like that of the non-nucleotide-linked sugar acetyltransferases, such as SAT and GAT. Class 2 enzymes, such as QdtC, WlbB, FdtD and AntD, lack the conserved histidine residue of the class 1 enzymes, and their reaction mechanisms likely proceed through substrate-assisted catalysis.

1.2.4 Non-L*β*H acyltransferases kinetic and chemical mechanisms

There are many non-L*β*H acyltransferases that are involved in a wide range of cellular functions, including posttranslational modification of nucleosomes, acetylation of antibiotics, fatty acid *β*-oxidation, methionine and acetylcholine biosynthesis amongst others. Further details of exemplars of these acyltransferases are presented in Table 6.

Table 6. Exemplars of non-L*β*H, hexapeptide acyltransferases

Acyltransferase [a]	acyl donor	acyl acceptor	Mechanism [b]	quaternary structure (pdb code)
Carnitine acetyltransferase (CRAT)[130,131]	Acyl CoA	Carnitine	Ternary complex	Monomer (1NM8)
Chloramphenicol acetyltransferase (CAT)[132,133]	Ac-CoA	Chloramphenicol	Ternary complex	Trimer (3CLA)
Choline acetyltransferase (ChAT)[134]	Ac-CoA	Choline	Ternary complex	Monomer (2FY2)
K(lysine) acetyltransferase 2B (KAT2B/PCAF)[135,136]	Ac-CoA	Lysine	Ternary complex	Monomer (1CAM0)
Serotonin *N*-acetyltransferase (SNAT/AANAT)[137]	Ac-CoA	Serotonin	Ternary complex	Monomer (1CJW)
Homoserine *O*-acetyltransferase (HTA)[138,139]	Ac-CoA	L-homoserine	Ping-pong	Monomer (2B61)
Homoserine *O*-succinyltransferase (HST)[140,141]	Succinyl-CoA	L-homoserine	Ping-pong	Dimer (2VDJ)
Arylamine *N*-acetyltransferase (NAT)[76,142,143]	Ac-CoA	Xenobiotic acceptors	Ping-pong	Dimer (4BGF)
Lecithin-cholesterol acyltransferase (LCAT)[144,145]	PC [c]	Cholesterol	Ping-pong	Monomer (5TXF)

[a] Citations comprise of structural biology, in addition kinetic and chemical mechanism studies. [b] Ternary complex mechanism refers to sequential bi-bi mechanisms, which can either be Theorell-Chance, ordered or random. [c] Phosphatidylcholine (PC).

The structures of non-L*β*H acyltransferases in Table 6 are diverse, spanning many different structural motifs. Two acyltransferases presented belong to the GNAT superfamily, namely PCAF and SNAT, which both catalyse acetyl transfer to acceptor substrates by an ordered sequential kinetic mechanism. During formation of the pre-catalytic, ternary complex, Ac-CoA binds to the free enzyme first, followed by the acceptor substrate, which is either *N*-terminal tail of a histone or serotonin, respectively. SNAT catalyses the penultimate step in the biosynthesis of melatonin, which is a hormone known to be involved in circadian rhythms, and subsequently human mood and behaviour. SNAT chemical mechanism elucidation (using the sheep enzyme), achieved using pH-rate, bisubstrate inhibition and SDM studies, identified His120, His122 and Trp168 active site residues as critical for catalysis. A general base mechanism was proposed, involving either His120 or His122 acting as the catalytic residue and Trp168 as the general acid[137]. The human PCAF chemical mechanism was proposed to proceed by general base catalysis, with Glu570 abstracting a proton from the ε-amino of Lys14 of histone H3, facilitating nucleophilic attack[135].

There are several non-L*β*H acyltransferases that catalyse acyl transfer using a ping-pong kinetic mechanism, which involves formation of an acyl-enzyme intermediate, prior to transfer to the acceptor substrate. The HTA protein structure comprises a two-domain organisation, which identifies the enzyme as a member of *α*/*β*-hydrolase superfamily. The canonical fold consists of an eight-stranded mainly parallel *β*-sheet, in which the second strand is oriented in the antiparallel direction, surrounded by a total of six *α*-helices. The chemical mechanism was determined using the *H. influenzae* HTA enzyme, with Ser143 identified as the catalytic nucleophile that attacks the thioester bond of Ac-CoA. A catalytic triad, reminiscent of serine proteases, was identified with His337 acting as the nucleophile-activating base and Asp304 taking the role of the orienting and polarising acid[138].

Carnitine and Chloramphenicol acetyltransferase are two-domain polypeptides that share a similar backbone fold, which is comprised of a *N*-terminal domain of an eight stranded *β*-sheet surrounded by eight *α* -helices, followed by a C-terminal domain of a six stranded *β*-sheet that is covered on one side by eleven *α* -helices. The quaternary structure of CAT is a homo-trimer, with three active sites formed at the interface of a pair of monomers. CRAT exists as a monomer, with the two domains

making numerous interactions that are akin to the monomer interactions present in the CAT homo-trimer[146]. The CAT and CRAT enzymes bind CoA-SH in a similar confirmation to the L*β*H acyltransferases, with a Lys positioned to interact with the 2'-phosphate of CoA-SH, and small residues, two Ser and an Asp packed against the pantothenic acid. A His residue is found in both CAT and CRAT active sites (His195 and His343, respectively), which is within hydrogen-bonding distance to an Asp or Glu residue (D199 and E347, respectively), forming a putative catalytic-dyad[147]. Sequential kinetic mechanisms are proposed for both non-L*β*H acyltransferases, with a rapid equilibrium kinetic mechanism proposed for both enzymes, with either a random or ordered mechanisms postulated for CAT_{III} (type III enzyme[133]) and mouse CRAT, respectively.

A conserved chemical mechanism is evident for all acyltransferase enzymes that catalyse acyl transfer by a ternary complex, namely a general base, general acid mechanism, utilising an active site His. The catalytic His residue is generally in a catalytic dyad with an acidic residue, which acts to polarise and orientate the imidazole side chain of the general base.

1.2.5 Antibiotics, resistance and acyltransferases

Antibiotic was first used as a noun in 1941 by Selman Waksman, the scientist responsible for identifying natural product drugs, such as actinomycin, streptomycin, neomycin, amongst others. Waksman received a Nobel Prize in 1953 for the discovery of streptomycin[148] (Figure 7).

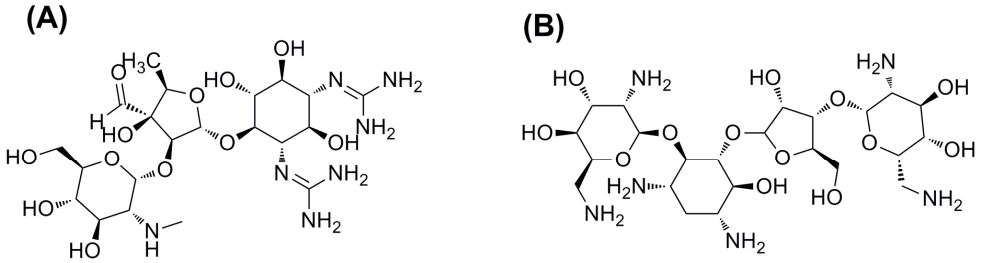


Figure 7. Streptomycin and Neomycin.

The aminoglycoside antibiotics that were discovered in the 1940s by Selman Waksman. (A) Streptomycin was discovered in 1943. (B) Neomycin was discovered in 1949.

Antibiotics are generically defined as any class of organic molecule that inhibits or kills microbes by specific interactions with bacterial targets, without any consideration of the source of the compound or class. This definition encompasses purely synthetic therapeutics, as these molecules specifically interact with receptors and provoke specific cell responses, in the same manner as inhibitors derived from natural products[149]. Natural products are defined as organic chemical compounds, invariably produced by primary and secondary metabolism pathways, in living organisms. The natural product definition, is generally further restricted to encompass only secondary metabolites, which are not essential for survival of organisms but endow an evolutionary advantage to the producer, as they have been optimised to be cytotoxic to prey, predators and competing organisms[150].

Synthetically derived antibiotics are also susceptible to the same biochemical mechanisms of cross-resistance in pathogens, as naturally derived molecules. Examples of synthetically derived antibiotics where resistance in pathogenic bacteria has appeared, include the fluoroquinolones (FQs), sulfonamides, and trimethoprim[151].

Historically, most clinically relevant antibiotics originate from soil-dwelling actinomycetes, including *β*-lactams, tetracyclines, rifamycins, aminoglycosides, macrolides, and glycopeptides[152]. Bacteria that produce antibiotics, such as the actinomycetes, harbour resistance elements for self-protection that are often clustered in antibiotic biosynthetic operons. Resistance elements, or genes orthologous to these, have been identified on mobile genetic elements, in antibiotic resistance pathogens in clinical settings. The molecular mechanisms of antibiotic resistance are varied and can occur by one or more of the following processes; modification of a drug target, molecular bypass, active efflux, decreased uptake and chemical modification of the compound. Target modification is often typified by the inclusion of a point mutation in key genes, which results in rapid resistance where the changes to the target have minimal impact on microbe fitness. Molecular bypass is where microbes have evolved mechanisms that avoid antibiotic action by utilising alternate pathways that are not antibiotic sensitive. Efflux involves the active removal of antibiotics from within the microbe using membrane-spanning efflux proteins. Finally, chemical modification is achieved by specifically adapted or evolved enzymes that catalyse the inactivation of antibiotics. The evolution of highly specific and efficient catalysts is evidence of strong and sustained selective pressure of

antibiotic action[151]. There are numerous examples of enzyme types that inactivate antibiotics, including hydrolases, kinases, glycosyltransferases, ribosyltransferases, adenylytransferases and acetyltransferases.

There are several examples of antibiotic inactivation by acetyltransferases, including aminoglycoside antibiotic *N*-acetyltransferases (AAC), chloramphenicol acetyltransferases (CAT) and xenobiotic acetyltransferases (XAT). The AACs are members of the GNAT superfamily and are known to acetylate and inactivate aminoglycosides, including Kanamycin, Amikacin, Neomycin and Streptomycin[153]. Recent evolution of the function of a subset of AACs has endowed these acetyltransferases with the ability to acetylate the synthetic fluoroquinolone class of antibiotics, such as ciprofloxacin[154].

Chloramphenicol (CAM) and Virginiamycin are two natural product antibiotics that have a molecular mechanism that is inhibition of the bacterial ribosome. CAM is a broad-spectrum antibiotic that was first isolated form *Streptomyces venezulae* in 1948[155]. CAM, despite having a relatively high level of toxicity, was used to treat many infections in the late 20th century[156]. CAM is still widely used in many developing countries because of its broad spectrum of activity and affordability, while in the developed world it is used to treat ophthalmic infections and as a last resort in life-threatening brain infections. The molecular target of CAM is the 50S subunit of the bacterial ribosome, where it binds the A-site and prevents the binding of the amino-acyl moiety of the A-site tRNA, leading to inhibition of protein synthesis[157].

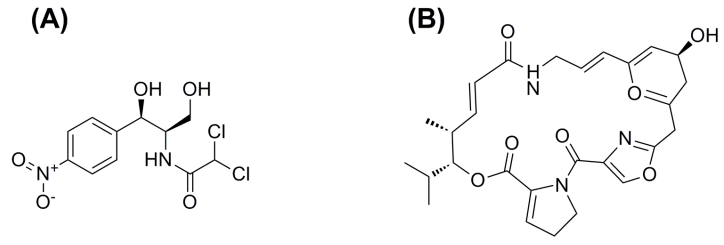


Figure 8. Natural product antibiotics that are acetylated by bacterial acetyltransferases.

The antibiotics in this figure are both deactivated via acetylation by bacterial acetyltransferases. (A) Chloramphenicol (CAM) and (B) Virginiamycin M_1, a streptogramin-class antibiotic.

Virginiamycin, is a member of the streptogramin, natural product antibiotic family that was first isolated from *Streptomyces pristinaspiralis* and *S. virginiae* in 1953[158]. The streptogramin antibiotics can be divided into two chemically distinct classes of cyclic molecules, type A and B, which are chemically unrelated but structurally similar. Streptogramins are used for the treatment of serious Gram-positive pathogens, including methicillin-resistant *S. aureus* and vancomycin-resistant *E. faecium*[112]. The streptogramins molecular mechanism is binding to the 50S subunit of the bacterial ribosome and interfering with peptidyltransferase activity. The type A and B streptogramins bind to separate but overlapping binding sites in the ribosomal P site. Binding of the type A streptogramins induces a ribosomal conformation change that significantly increases the binding affinity of the type B streptogramins[159]. Type A streptogramins inhibit protein synthesis by interfering with substrate attachment to both A and P sites of the A and P sites of the 50S ribosome. Type B streptogramins block peptide bond synthesis during elongation by causing incorrect positioning of the peptidyl tRNA at the P site[160].

Resistance to multiple antibiotics was first observed in Japan in the 1950s, including a resistance determinant for CAM on mobile genetic elements, in this case plasmids responsible for transferable phenotype. The molecular mechanism for this CAM resistance was shown to be an enzymatic inactivation by *O*-acetylation of the C3 primary hydroxyl of CAM, which subsequently prevented the acetylated product from binding to bacterial ribosomes. CAT was identified as the enzyme responsible for CAM acetylation and subsequently the acetyltransferase was identified as being widespread and well conserved, among the bacterial genera. The type III CAT enzyme is the most active of the known CAT variants and the determined steady-state parameters highlight that this enzyme has evolved to be an efficient catalyst of CAM.

In addition to the CAT_{III} class of chloramphenicol acetyltransferases, a new class of acyltransferases was identified that were able to modify CAM in 1991 from *Agrobacterium tumefaciens*[161]*.* The new class of CAT, classified as Xenobiotic acetyltransferases (XAT), share no homology with the previously characterised CAT_{III} enzymes but can catalyse *O*-acetylation of the C3 and C1 primary hydroxyls of CAM. An early exemplar of the XAT enzymes, the *satA* gene product from *Enterococcus faecium,* was shown to inactivate, by acetylation, the streptogramin group A antibiotics, including virginiamycin M1, along with low level acetylation of

chloramphenicol[162]. Subsequent studies identified several XAT acetyltransferases from a diverse selection of bacteria, including *E. coli*, *P. aeruginosa*, *S. aureus*, all of which shared the same L*β*H structural motif[82,163]. Streptogramin resistance is now widespread and has arisen from multiple mechanisms, with chemical modification through acetylation one of the most pertinent, have reduced the effectiveness of these compounds. This antibiotic resistance caused by chemical modification of inhibitors via acetylation, highlights that the deployment of L*β*H acyltransferases is an effective bacterial survival mechanism.

1.3 Peptidoglycan biosynthesis and GlmU

1.3.1 Peptidoglycan biosynthesis in bacteria

The cell wall is a unique macromolecule common to both Gram-positive and Gram-negative bacteria but the biochemical and structural composition vary significantly between bacterial species[164]. Peptidoglycan (or murein), as previously discussed is composed of linear glycan strands that are cross-linked by short peptides. The repeating glycan strands are made up of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) or *N*-glycolylmuramic acid (UDP-MurNGlyc) residues, which are linked by β-(1,4) bonds (Figure 9). The peptidyl component of *M. tuberculosis* peptidoglycan is formed by substitution of the D-lactoyl group of each MurNAc residue with a peptide stem, which is often composed of L-Ala-γ-D-Glu-*meso*-DAP (or L-Lys)-D-Ala-D-Ala (*meso*-DAP, 2,6-diaminopimelic acid)[165]. The *Mycobacterium* peptidoglycan can be cross-linked, typically between corresponding tetrapeptide side chains, comprising of either “3,3” (between *meso*-DAP residues) or “4,3” (D-Ala to *meso*-DAP residues). The *Mycobacterium* species (*spp.)* peptidoglycan is highly cross-linked (70-80%) in comparison to *E. coli* (30-50%), which provides added structural integrity (Figure 10)[166].

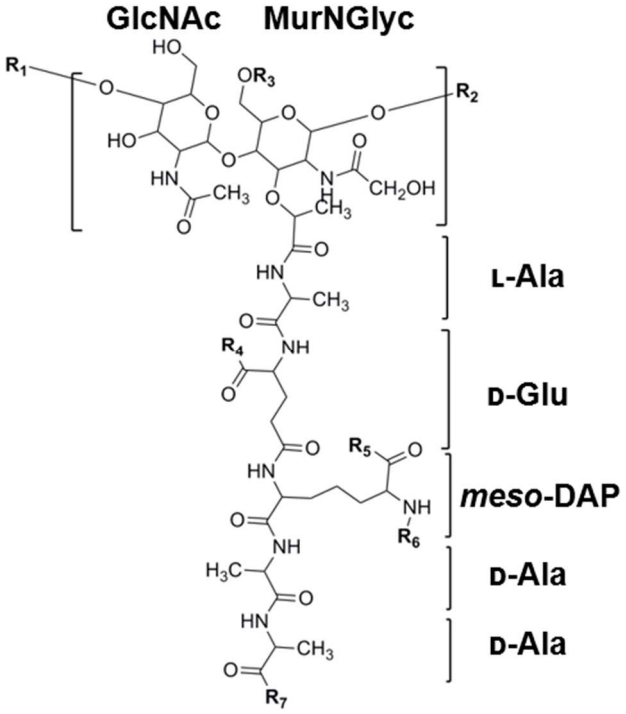


Figure 9. Structure of a representative peptidoglycan monomer.

R_1, *N*-glycolymuramic acid residue of another monomer; R_2, *N*-acetylglucosamine residue of another monomer; R_3, H or the linker unit of arabinogalactan; R_4, OH, NH_2 or Gly; R_5, OH or NH_2; R_6, H, or cross-linked to penultimate D-Ala or the D-centre of another meso-DAP residue; R_7, OH or NH_2.

The biosynthesis of peptidoglycan is a complex, multi-enzyme process that involves approximately 20 reactions[167]. The initial biosynthesis steps occur in the cytoplasm, which predominantly encompasses the synthesis of the nucleotide precursors of peptidoglycan. The peptidoglycan precursor is then acted upon at the inner side of the cytoplasmic membrane to synthesise lipid-linked intermediates, followed by polymerisation reactions on the extracellular surface of the phospholipid bilayer[168]. The cytoplasmic steps of peptidoglycan biosynthesis can be broadly divided into four processes, which include formation of UDP-*N*-acetylglucosamine (UDP-GlcNAc) from fructose 6-phosphate (Fruf-6P), followed by generation of UDP-*N*-acetylmuramic acid (UDP-MurNAc) and subsequent modification to add the pentapeptide that is assembled from the side pathway synthesis products D-Glu and D-alanyl-D-alanine[169].

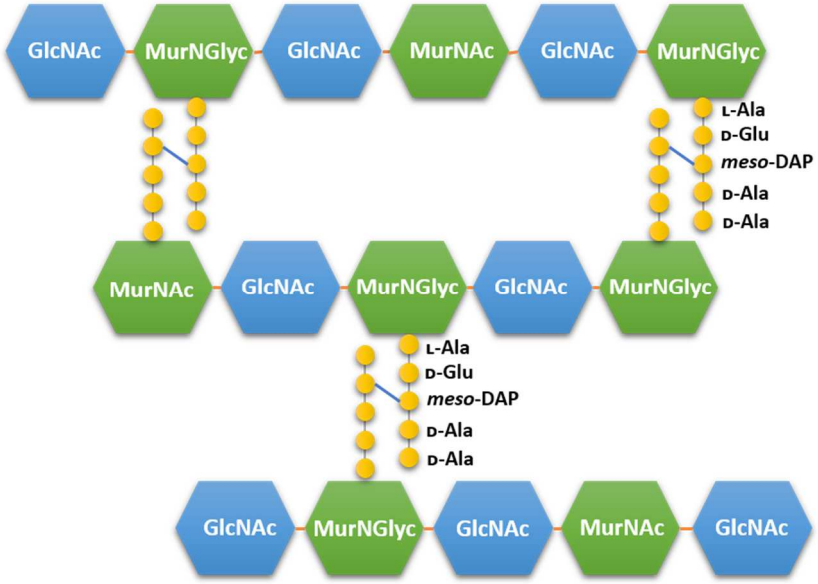


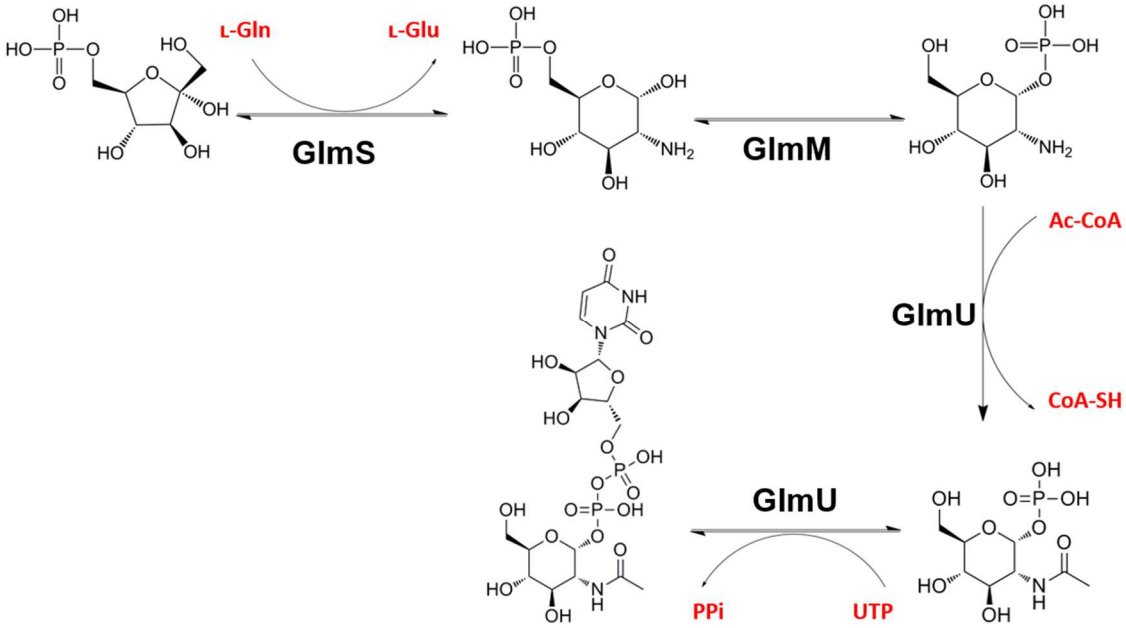
Figure 10. Representative schematic of the arrangement of *Mycobacterium* peptidoglycan.

Peptidoglycan biosynthesis begins with the generation of UDP-GlcNAc from Fruf-6P in a four reaction, three enzyme biosynthetic pathway (Scheme 3). The three enzymes involved in the UDP-GlcNAc biosynthetic pathway are glucosamine 6-phosphate (GlcN-6P) synthase (GlmS), phosphoglucosamine mutase (GlmM) and the bifunctional *N*-acetyl-glucosamine-1-phosphate (GlcNAc-1P) acetyltransferase uridylyltransferase (GlmU).

GlmS is a glutamine-dependent amidotransferase that catalyses the formation of D-glucosamine 6-phosphate (GlcN-6P) from L-glutamine (L-Gln) and Fruf-6P[170]. GlmS is an essential enzyme comprised of two structurally and functionally distinct domains, which are an *N*-terminal glutaminase domain and a *C*-terminal isomerase domain. The glutaminase domain promotes glutamine hydrolysis into glutamate and ammonia, while the isomerase domain binds the previously formed products to catalyse the conversions of Fruf-6P into GlcN-6P[171]. The catalytically active enzyme is a homodimer and the kinetic mechanism is a sequential ordered bi-bi mechanism where Fruf-6P binds first followed by L-glutamine[172]. The *M. tuberculosis* GlmS has not been studied but sequence alignments and homology modelling indicate that the polypeptides fold in a similar manner and the catalytic residues identified by studying the *E. coli* orthologue are conserved[173].

GlmM catalyses the interconversion of D-glucosamine 6-phosphate to D-glucosamine 1-phosphate[174]. The *E. coli* orthologue was the first to be identified, purified to near homogeneity and subsequently characterised kinetically. The structure of GlmM was determined for both the *Bacillus anthracis* (pdb 3PDK) and *S. aureus* (pdb 6GYZ) and the tertiary structure was arranged as four structural domains, with domains 1 to 3 consisting of a mixed α/β core and domain 4 comprising a three stranded antiparallel β-sheet, flanked by two α-helices[175].

GlmU is a bifunctional enzyme that catalyses the *N*-acetylation of GlcN-1P and the uridylylation of GlcNAc-1P to form UDP-GlcNAc. GlmU will be described in more depth in section 1.3.2.



Scheme 3. Formation of UDP-GlcNAc from Fruf-6P and L-Gln.

The generation of UDP-MurNAc and subsequent modification with the pentapeptide, which is synthesised in a sequential pathway catalysed by the Mur ligases A to F. The first two steps of UDP-MurNAc are catalysed by MurA, a UDP-N-acetylglucosamine 1-carboxyvinyltransferase and MurB, a UDP-N-acetylenolpyruvoylglucosamine reductase. MurA catalyses the addition of the enoylpyruvyl moiety of phosphoenoyl pyruvate to UDP-GlcNAc, followed by MurB reduction of UDP-GlcNAc-*enol*-pyruvate to a lactyl ether of UDP-MurNAc using an equimolar amount of NADPH and a solvent derived proton. The next step of peptidoglycan biosynthesis, formation of UDP-MurNGlyc, is rare and only occurs in

the *Mycobacterium spp.* and a small number of closely related genera of bacteria. The *M. tuberculosis* enzyme NamH (Rv3808), a UDP-N-acetylmuramic acid hydroxylase, catalyses the formation of UDP-MurNGlyc, by hydroxylation of the methyl group of the MurNAc moiety of UDP-MurNAc, providing both types of UDP-muramyl substrates[176]. *Mycobacterium spp.* peptidoglycan is comprised of a mixture of both UDP-MurNAc and UDP-MurNGlyc, which confers an increase in intrinsic strength, by potentially alleviating the susceptibility of UDP-MurNAc to lysozyme, in addition to an elevation in the number of potential hydrogen bonding sites[176].

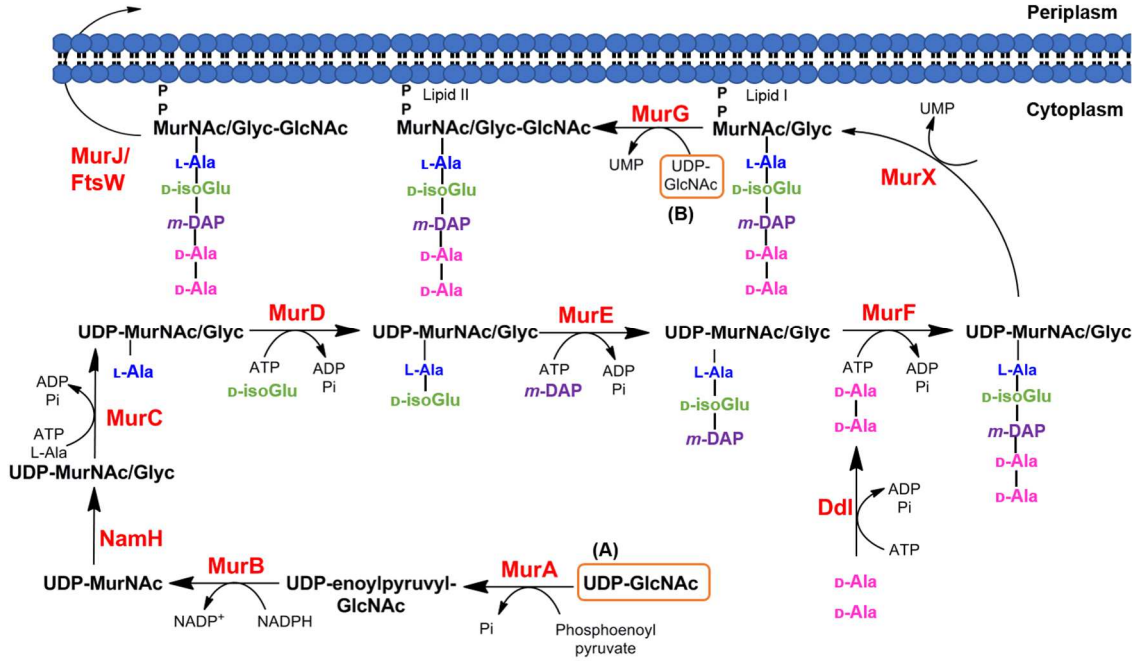


Figure 11. A schematic of the cytoplasmic steps of *M. tuberculosis* peptidoglycan biosynthesis.

Utilisation of UDP-GlcNAc as a substrate by (A) MurA and (B) MurG highlighted. Figure inspired by similar images from a number of publications, including Abrahams and Besra[177], Typas, *et al.*[178] and Chellat, *et al.*[179].

The UDP-MurNAc and UDP-MurNGlyc intermediates are then processed through the peptidoglycan biosynthetic pathway beginning with UDP-*N*-acetylmuramoyl:L-alanine ligation by MurC. This is followed by D-isoglutamate ligation by MurD, *meso*-DAP addition by MurE, and finally D-alanyl-D-alanine ligation by MurF (Figure 11)[168]. The final product of this ligation cascade is known as Park's nucleotide[180], a muramylpentapeptide: UDP-MurNAc/Glyc-L-Ala- γ-D-Glu -*meso*-DAP-D-Ala-D-Ala[181]. There have been numerous reviews of the ATP-dependent Mur ligase pathway in recent years[168,182-184]; and there is increasing information regarding the *M.*

tuberculosis enzymes. A recent study reconstituted the *M. tuberculosis* MurA-F pathway *in vitro*[185], in addition MurA (Rv1315), MurB (pdb 5JZX, Rv0482)[186], MurC (Rv2151c)[187], MurD (Rv2155c)[188], MurE (pdb 2WTZ, Rv2158c)[189] and MurF (Rv2157c) [190] have been biochemically characterised. The Mur ligases C-F have different amino acid specificities; however, the four enzymes share common properties, including kinetic mechanism, six invariant residues that are characteristic of the Mur ligases, a preserved ATP-binding sequence and conserved three-dimensional structure[168].

The final cytoplasmic enzyme involved in peptidoglycan biosynthesis is D-Ala-D-Ala ligase (Ddl), which catalyses the ATP-dependent formation of the D-Ala-D-Ala dipeptide. Prior to synthesis of the D-Ala-D-Ala dipeptide, the racemisation of D-Ala from L-Ala, is catalysed by alanine racemase (Alr), which is a pyridoxal 5-phosphate-dependent enzyme[191,192]. The formation of D-Ala-D-Ala by Ddl provides MurF with substrate for the final cytoplasmic ligation reaction to form Park's nucleotide. The *M. tuberculosis* Ddl (Rv2981c) has been structurally characterised (pdb 3LWB)[193] and the kinetic mechanism, which is an ordered ter-ter mechanism, has been elucidated[191].

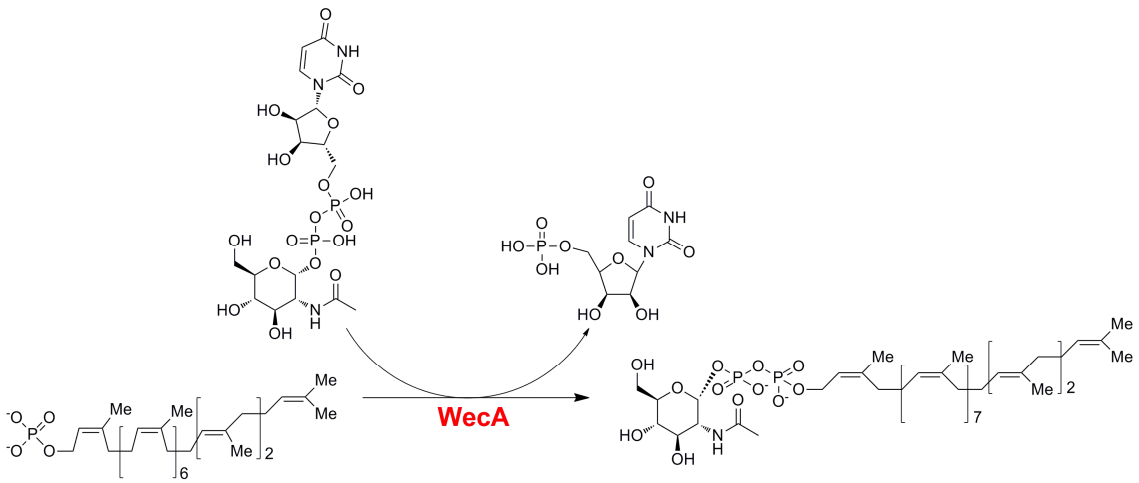
The first membrane-associated peptidoglycan precursor is generated by the translocation of Park's nucleotide to decaprenyl phosphate (C_{50}-P), catalysed by MurX, Phospho-*N*-acetyl-muramoyl-pentappeptide-transferase, (also known as MraY)[194], forming Lipid I. The final intracellular step of peptidoglycan synthesis is catalysed by the glycosyltransferase, MurG, which forms a β-(1,4) linkage between GlcNAc (from UDP-GlcNAc) and MurNAc of MurNGlyc of Lipid I, leading to the generation of Lipid II, the monomeric building block of peptidoglycan[195]. To date the biochemical characterisation of the *M. tuberculosis* MurX (Rv2156c) and MurG (Rv2153c) enzymes has not been documented.

1.3.2 Bifunctional acetyltransferase/uridylyltransferase, GlmU

The *Mycobacterium* biosynthetic pathway of UDP-GlcNAc is critical for the formation of two key components of the cell wall. UDP-GlcNAc is the core cytoplasmic precursor of peptidoglycan, incorporated as one of the two amino sugars that make up the disaccharide subunit of the linear glycan strand[165]. Addition of UDP-GlcNAc to Lipid I to form Lipid II, the final monomeric unit of peptidoglycan, is catalysed by

MurG. Additionally, UDP-GlcNAc is utilised as the sole, initial donor for generation of the second amino sugar of the peptidoglycan disaccharide subunit, UDP-MurNAc[196]. To form the UDP-MurNAc precursor UDP-enoylpyruvyl-GlcNAc, MurA catalyses enoylpyruvyl transfer from phospho-enoyl-pyruvate to UDP-GlcNAc.

In addition to the dual utilisation in the generation of peptidoglycan, UDP-GlcNAc is also the initiating donor of GlcNAc-1P for the formation of D-*N*-acetylglucosamine-L-rhamnose (D-*N*-GlcNAc-L-Rha), which is the linker unit between peptidoglycan and arabinogalactan[197]. The first step in the biosynthesis of D-*N*-GlcNAc-L-Rha is catalysed by GlcNAc-1-phosphate transferase (WecA, Rv1302), which generates decaprenyl phosphate-GlcNAc from UDP-GlcNAc and decaprenyl monophosphate (DecP) (Scheme 4)[198]. Interestingly, cell wall biosynthesis in Gram-negative bacteria, such as *E. coli* and *S. typhimurium*, requires UDP-GlcNAc, which is one of two essential precursors for Lipid A (or Endotoxin) synthesis. The initial reaction in the biosynthesis of Lipid A is the acylation of UDP-GlcNAc by LpxA, another L*β*H acyltransferase enzyme, using ACP thioesters as the donor substrate[199].



Scheme 4. UDP-GlcNAc utilisation in the first step of arabinogalactan biosynthesis.

WecA catalysed formation of decaprenyl phosphate-GlcNAc, the first membrane-anchored glycophospholipid in the arabinogalactan biosynthetic pathway.

The enzymes and the reactions involved in the biosynthetic pathway of UDP-GlcNAc were first postulated and tentatively identified by a number of groups between the period of 1968-70[200-202], however, it wasn't until 1988 until the first member of the cascade, GlmS, was purified and initially characterised[203]. Previous *E. coli* DNA sequencing studies in 1984, characterising the nucleotide sequence around *unc*

operon identified an open-reading frame (ORF) preceding the *glmS* gene that was initially designated *Eco*URF-1[204]. A gene from *Bacillus subtilis*, *tms*, was identified in 1989, with an ORF of 456 codons, resulting in a putative translation product with an expected molecular weight of 49.55 kDa[205]. The *tms* gene product was postulated to share 43% amino acid sequence similarity with the *Eco*URF-1 gene product and was later re-designated as *gcaD* and identified as a GlcNAc-1P uridylyltransferase[206]. Subsequent studies were carried out on the *E. coli Eco*URF-1 gene, in 1993, which found that a thermosensitive, conditional mutation led to inhibition of peptidoglycan synthesis, upon impairment of the gene synthesis. Additionally, weakening of *Eco*URF-1 gene synthesis at 43 °C caused alterations to *E. coli* morphology and ultimately cell lysis, which when analysed had 37% lower peptidoglycan content than normally growing bacteria. Further analysis of the lysates identified a large accumulation of GlcNAc-1P and reduction of the levels of the seven peptidoglycan nucleotide precursors. These peptidoglycan quantification data identified that the conditional mutation introduced blocked uridylylation of GlcNAc-1P. Overexpression of the *Eco*URF-1 gene, re-designated as *glmU,* in *E. coli* led to a large increase in GlcNAc-1P uridylyltransferase activity and the application of a two-chromatographic step protein purification strategy enabled isolation of homogeneous enzyme[207].

Further studies utilising the *E. coli glmU* gene product, in 1994, identified that GlmU, a 456-amino acid, 49.13 kDa protein, is a bifunctional enzyme that catalyses the preceding step of GlcN-1P acetyl transfer, in addition to GlcNAc-1P uridylylation. The kinetic parameters for both GlmU activities were characterised, in addition to initial identification that the acetyltransferase activity of the *E. coli* enzyme was sensitive to thiol-alkylating agents, suggesting the presence of a catalytically important Cys residue[208]. Additional characterisation of the *E. coli* GlmU kinetic mechanism elucidated the order of the bifunctional enzyme activities with the GlcN-1P acetyltransferase activity preceding the GlcNAc-1P uridylyltransferase function. Domain dissection experiments demonstrated that acetyltransferase activity was associated to the *C*-terminal domain of GlmU and the *N*-terminal domain was identified as the location of the uridylyltransferase active site[209,210].

The three-dimensional structure of a truncated version *E. coli* GlmU was initially determined in 1999 and was subsequently followed by the full-length version in 2001, both achieved by X-ray crystallography[211]. The GlmU tertiary structure is divided into two distinct domains located at the *N*- and *C*-terminals of the polypeptide and the

quaternary structure is comprised of a homotrimer, arranged around a threefold crystallographic axis of symmetry[212]. Subsequently, the structures of GlmU orthologues from several bacteria including; *Streptococcus pneumoniae* (PDB ID 1G95), *Haemophilus influenzae* (PDB ID 2V0H) and *M. tuberculosis* (PDB ID 3D98) have been determined[213-215]. These structures show a high degree of similarity to the *E. coli* orthologue and confirm that in all cases the monomer of GlmU is folded into two distinct domains. The *C*-terminal domain (residues 263-478) is formed by a L*β*H, which is similar to a number of acetyl- or acyltransferase enzymes, such as SAT, GAT, MAT, PaXAT, DapD, LpxA and LpxD, which make up the L*β*H superfamily[73,78,79,81,95,216]. As previously described in section 1.2.2, proteins belonging to the L*β*H superfamily have a structure comprising a parallel β-helix with repeating isoleucine-rich hexapeptide motifs and left-handed connections (Figure 12).

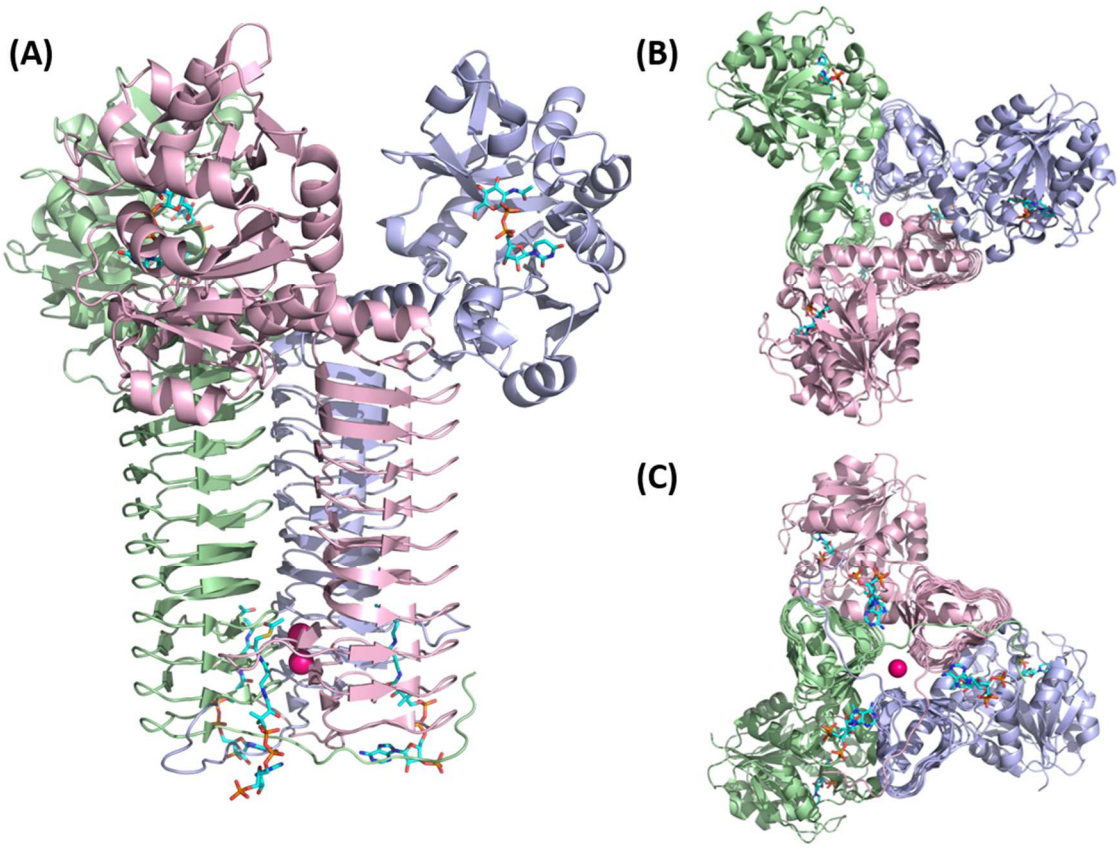


Figure 12. Structure of *E. coli* GlmU, a L*β*H acetyltransferase.

This figure depicts the *E. coli* GlmU (pdb 2OI5) quaternary structure. Each monomer is displayed as a cartoon representation of the secondary structure. (A) Side view of the trimer, showing two Mg^{2+} ions bound, (B) top view of the GlmU trimer and (C) representation of the bottom of the GlmU trimer.

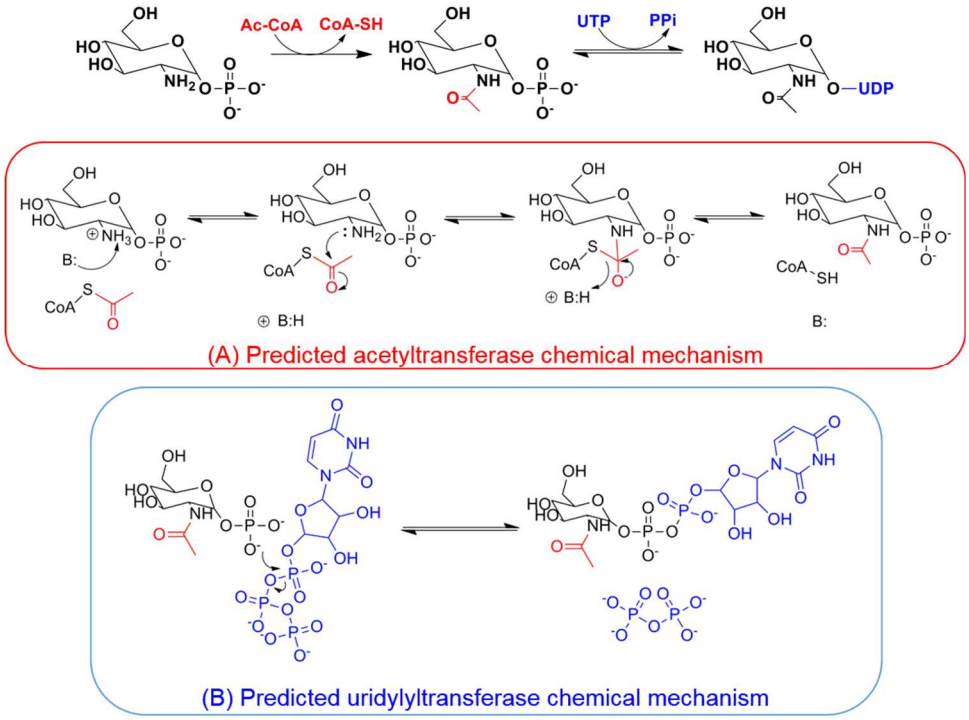
The *C*-terminal domain of *E. coli* GlmU is similar to the canonical L*β*H fold and is the longest example observed, comprising eleven complete or partial coils (Figure 12 (A)). The prismatic coil of the L*β*H domain is interrupted once between residues Thr384 and Thr395 by a 10-residue inserted loop that projects from the T3 turn of coil C8 into the inter-subunit space and adjacent to the thiol group of CoA-SH. The *E. coli* homotrimer has two Mg^{2+} ions bound on the three-fold axis, which are coordinated by the three copies, one from each monomer, of the side chain carboxylate of Asp406 (Figure 12 (A)). The key distinguishing feature of the GlmU L*β*H is the truncation of the T3 turn, which are shortened by one residue for five of the eleven turns. The T3 turn position of the L*β*H is typically where loops are inserted in hexapeptide enzymes, which provide a means by which these acyltransferases attain structural and functional diversity[217]. The acetyltransferase active site of GlmU, in a similar manner to other members of the L*β*H family, is located at a cleft between two monomers of the homotrimer, with interacting residues contributed from both polypeptides. GlmU is unique amongst the L*β*H acyltransferases in its use of all three subunits to form its acetyltransferase active site and in the interactions of its 19-residue *C*-terminal tail extension with the cofactors of two distinct acetyltransferase active sites (Figure 12 (C)). The GlmU orthologues that have been structurally characterised using X-ray crystallography are detailed in Table 7.

Table 7. Published GlmU structures in the Protein Data Bank (pdb)

GlmU orthologue	Number of structures	pdb code (s)	Ligands investigated
E. coli	6	4AA7, 3TWD, 2OI5, 2OI6, 2OI7, 1HV9	Anti-bacterial inhibitor (x2), UDP-GlcNAc, Ac-CoA, CoA-SH, desulpho-CoA, Mg^{2+}
H. influenzae	14	4KNR, 4KNX, 4KPX, 4KPZ, 4KQL, 4EIK, 2W0V, 2W0W, 2V0H, 2V0I, 2V0J, 2V0K, 2V0L, 2VD4	Anti-bacterial inhibitor (x7), UDP-GlcNAc, Ac-CoA, CoA-SH, UDP, Mg^{2+}
S. pneumoniae	7	4AAW, 4AC3,1HM0, 1HM8, 1HM9, 1G95, 1G97	Anti-bacterial inhibitor (x2), UDP-GlcNAc, Ac-CoA, UDP, Mg^{2+}, Ca^{2+}
M. tuberculosis	15	6GE9, 4K6R, 4G3P, 4G3Q, 4G3S, 4G87, 4HCQ, 3SPT, 3ST8, 3FOQ, 3DJ4, 3DK5, 3D8V, 3D98, 2QKX	UDP-GlcNAc, Ac-CoA, Glc-1P, GlcN-1P, GlcNAc-1P, ATP, CoA-SH, PP_i, Mg^{2+}, Co^{2+}
Yersinia pestis	2	4FCE, 3FWW	GlcN-1P, Mg^{2+}
A baumannii	1	5VMK	N.A.

The pyrophosphorylase *N*-terminal domain of *E. coli* GlmU was the first of its type to be crystallised and have its three-dimensional structure solved[218]. Sequence comparison of the *E. coli* GlmU *N*-terminal domain with other nucleotide sugar pyrophosphorylase domains, including ADP-Glc, UDP-Glc and GDP-Man, reveals similarities ranging from 12 - 24% and a close resemblance with the dinucleotide-binding Rossmann fold[211,219]. The overall structure of the *E. coli N*-terminal domain pyrophosphorylase domain is that of a seven-stranded mixed *β*-sheet surrounded by six *α*-helices and a two-stranded *β*-sheet sitting atop the larger mixed sheet (Figure 12 (B)). The active site pocket is bounded by two lobes, the first of which interacts with the nucleotide moiety of UDP-GlcNAc and the second that primarily interacts with the GlcNAc portion of the sugar nucleotide. Divalent metal ion binding in the

pyrophosphorylase active site, coordinated by the side chains of Asp105 and Asn227, induces conformational change, which moves the first lobe toward the UDP-GlcNAc. The metal ion is also coordinated by two well-ordered H_2O molecule and the *α*– and *β*-phosphates of UDP-GlcNAc. A number of divalent metal ions, including Mg^{2+}, Co^{2+}, Ca^{2+} and Mn^{2+} have been shown to stimulate pyrophosphorylase activity of GlmU and the aforementioned nucleotide sugar homologues[212]. The first lobe comprises a mobile loop that has a sequence of L-X_2-G-X-G-T-X-M-X_4-P-K, which is a conserved pyrophosphorylase sequence fingerprint, between residue Leu11 – Lys25[212].



Scheme 5. GlmU catalysed formation of UDP-GlcNAc.

(A) Predicted GlmU acetyltransferase activity chemical mechanism. (B) Predicted GlmU uridylyltransferase activity chemical mechanism.

The catalytic mechanism of both of activities of bifunctional GlmU, from any bacterial species, has been inferred from structural biology studies, rather than kinetic studies, utilising classical enzymology techniques, such as initial velocity and inhibitor pattern studies, pH-rate profiles and SKIEs. The inferred acetyltransferase catalytic mechanism from X-ray crystallography studies, in addition to mutational analysis, applying SDM, is that the active site contains a catalytic triad comprised of a

conserved His residue, along with either an acidic Glu or Asp and a Ser (Scheme 5). The first substrate proposed to bind to GlmU is Ac-CoA, which is then followed by GlcN-1P. The acidic Glu or Asp (Glu349, *E. coli* GlmU) residue is proposed to orientate the His imidazole (His363, *E. coli* GlmU) in a catalytically favourable position, in addition to increasing the basicity of the sidechain. The Nε2 of the His imidazole is proposed to act as a general base, which deprotonates the 2-amino group of GlcN-1P, enabling subsequent nucleophilic attack on the Ac-CoA thioester. After nucleophilic attack, the resulting tetrahedral oxyanion intermediate is proposed to be stabilised by either the main-chain amide of either an Ala or the hydroxyl side chain of a Ser (Ala380 or Ser405, *E. coli* GlmU)[220]. However, despite the paucity of kinetic catalytic mechanism studies, steady-state kinetic parameters have been determined for several GlmU orthologues and are detailed in Table 8.

Table 8. Published GlmU steady-state kinetic parameters [a]

GlmU orthologue	$k_{cat,}$ acetyltransferase (s^{-1})	$k_{cat,}$ uridylyltransferase (s^{-1})	$K_{m,}$ Ac-CoA (mM)	$K_{m,}$ GlcN-1P (mM)	$K_{m,}$ UTP (mM)	$K_{m,}$ GlcNAc-1P (mM)
E. coli[209]	77	21	0.32	0.25	0.017	0.018
E. coli[210]	1,350	330	0.2	0.3	0.1	0.07
H. influenzae[221]	N.D.	N.D.	0.39	0.29	N.D.	N.D.
B. subtilis[222]	0.81 ± 0.01	N.D.	1.82 ± 0.3	0.33 ± 0.01	N.D.	N.D.
M. tuberculosis[223]	82.4 ± 3.2	3.1 ± 0.51	0.22 ± 0.07	0.061 ± 0.005	0.024 ± 0.002	0.044 ± 0.005
Y. pestis[224]	853 ± 69	95 ± 8	0.52 ± 0.1	0.3 ± 0.09	0.095 ± 0.027	0.061 ± 0.017

[a] Values collated from individual manuscripts and www.BRENDA-enzymes.org.

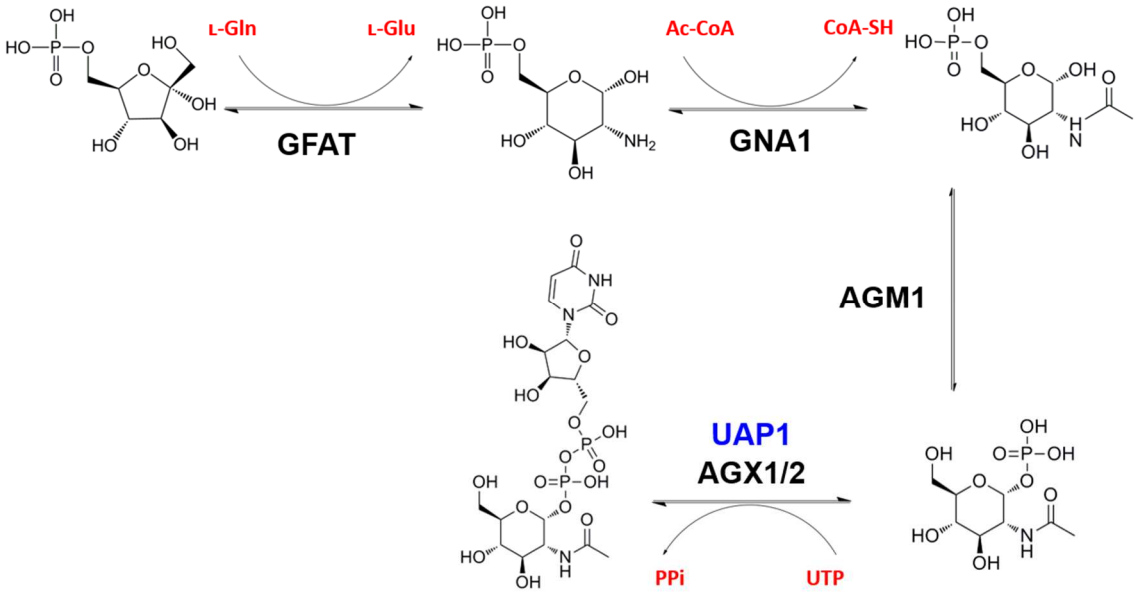
The inferred catalytic mechanism of GlmU catalysed uridylyl transfer has been extensively studied for several bacterial orthologues, primarily for the *H. influenzae* enzyme, using structural biology and SDM studies. The proposed condensation of UTP and GlcNAc-1P is thought to follow an ordered sequential mechanism (Scheme 5). The first substrate proposed to bind to the apoenzyme is UTP, which conserves the active site in the open conformation. The binding of UTP is followed by GlcNAc-1P, which is thought to elicit a conformation change that enables an environment for catalysis to occur, upon binding of Mg^{2+} coordinated by Asp and Asn residues (Asp102 or Asn227, *H. influenzae* GlmU). The phosphate group of GlcNAc-1P is then proposed to attack the phosphorous atom of the *α*-phosphate of UTP, which is stabilised by Arg and Lys residues (Arg15 or Lys22, *H. influenzae* GlmU), releasing PP_i and inversing the configuration of the *α*-phosphate[225].

The selective advantage to bacteria, conferred by having both acetyltransferase and pyrophosphorylase activities localised on the same polypeptide has yet to be elucidated[210]. In particular, there is no apparent requirement for a common regulation of the two activities at the level of either transcription or translation. The *glmU* gene is co-transcribed with *glmS* in *E. coli*, and has been reported to be expressed at a high constitutive level regardless of growth conditions[226]. In addition, studies into the reaction mechanisms of GlmU, utilising radiolabelled substrates, it was demonstrated that GlcNAc-1P was released by the enzyme before being used as a substrate for the second activity. In the same study, concentrations of GlcNAc-1P were observed to accumulate after acetyl transfer and prior to uridylyl transfer[209].

1.3.3 UDP-GlcNAc biosynthesis in eukaryotes

In eukaryotes, the pathway for UDP-GlcNAc biosynthesis is significantly different as acetyl transfer occurs on GlcN-6P and not GlcN-1P and, most importantly, acetyltransferase and uridylyltransferase activities are carried out by two distinct monofunctional enzymes (Scheme 6). In mammalian cells, the GlmS equivalent, known as, glutamine: fructose-6-P amidotransferase (GFAT), is an insulin-regulated enzyme that controls the flux of glucose into the hexosamine pathway[227]. GFAT is 280 kDa in size, and it is composed of four subunits; it belongs to the Ntn-amidotransferase family and possesses a Cys1 residue, as does the *E. coli* GlmS[228]. GFAT is subject to allosteric regulation by UDP-GlcNAc and can also be regulated,

via product inhibition, by GlcN-6P. In mammals, the pyrophosphorylases that condense UTP and GlcNAc-1P, are two isoforms of the same enzyme, known as AGX1 (pdf 1JV1) and AGX2 (pdf 1JVD)[229]. These enzymes are comprised of a large central core, flanked by two smaller extra domains, referred to as the *N*- and *C*-terminal domains. The central core domain comprises *α/β* structure arranged in a domain resembling the nucleotide binding Rossmann fold. The C-terminal domain of both enzymes contribute the majority of the interactions in formation of a catalytically active dimer. The AGX fold shares significant homology with the pyrophosphorylases, including GlmU, as well as Glc-1P thymidylyltransferase (RmlA) and CAMP-acylneuraminate (CAMP-NeuAc) synthetase[230]. The homologies between these enzymes are mainly centred on the Rossmann fold-like domains of the aforementioned enzymes, however, there are similarities beyond this domain. The long helix at the start of the *C*-terminal domain of AGX1/2 corresponds to the long helical arm that connects *N*- and *C*-terminal domains of GlmU[229]. In humans, *N*-acetylation of GlcN-6P to form GlcNAc-6P is carried out by the enzyme GlcN-6P *N*-acetyltransferase 1 (GNA1), which is an amino sugar *N*-acetyltransferase member of the GNAT superfamily[231]. The human enzyme involved in the conversion of GlcNAc-6P to GlcNAc-1P is a phosphoglucomutases, *N*-acetylglucosamine-phosphate mutase (either AGM1 or PGM3)[232,233].



Scheme 6. Formation of UDP-GlcNAc from Fruf-6P and L-Gln, in eukaryotes.

The biosynthetic pathway of UDP-GlcNAc formation in eukaryotes. The main difference between mammals and yeast is the exchange of AGX1/2 for UAP1.

In yeast, such as *Saccharomyces cerevisiae* and *Candida albicans*, the hexosamine metabolism has also been well studied and the four different enzymes involved in this pathway have been shown to be essential for cell viability[234]. The first reaction that leads to the formation of GlcN-6P from Fruf-6P, is catalysed by GFA1, which is a homo-tetramer of 80 kDa[235,236]. The second step is the *N*-acetylation of GlcN-6P by the GNA1 acetyltransferase to yield GlcNAc-6P[237]. The GNA1 is a member of the GNAT superfamily of acetyltransferases and has a three-dimensional structure that is comprised of a dimer of two identical subunits (pdf 1I21)[238]. The sequence similarity of GNA1 between *S. cerevisiae* and *C. albicans* is 44% and there is limited sequence similarity to GlmU[234]. The next step of UDP-GlcNAc biosynthesis is the isomerisation of GlcNAc-6P to GlcNAc-1P by AGM1, which is a 60 kDa GlcNAc phosphate mutase enzyme (pdb 2DKC)[239]. The final step in the UDP-GlcNAc biosynthetic pathways is the uridylylation of GlcNAc-1P by the enzyme UDP-*N*-acetylglucosamine-pyrophosphorylase (UAP1)[240]. The crystal structure of the *C. albicans* enzyme has been solved recently (pdb 2YQC)[241], which revealed a high degree of homology with the human and *Mus musculus* homologues.

1.3.4 *M. tuberculosis* GlmU

Sequencing of the *M. tuberculosis* H37Rv genome was completed in 1998, which allowed initial classification and annotation of the 4000 genes identified[242], followed by re-annotation in 2002[243]. GlmU was initially identified as being encoded by the *M. tuberculosis* Rv1018c gene and was subsequently cloned, expressed and purified in *E. coli* BL21(DE3) cells[244]. The purified protein from the aforementioned study was characterised with both acetyltransferase and uridylyltransferase activities identified. The three-dimensional structure of *M. tuberculosis* GlmU was determined in 2009, revealing a homo-trimer that shared a high degree of similarity with the previously determined structures for *S. pneumoniae*, *E. coli* and *H. influenzae* GlmU (Figure 13). To successfully crystallise the GlmU polypeptide, it has previously been speculated that it was necessary to co-crystallise the enzyme with either UDP-GlcNAc or Ac-CoA, which are believed to stabilise the *N*- and *C*-terminal domains, respectively[245]. The quaternary structure of GlmU in solution was confirmed as a homo-trimer utilising dynamic light scattering and analytical gel filtration experiments.

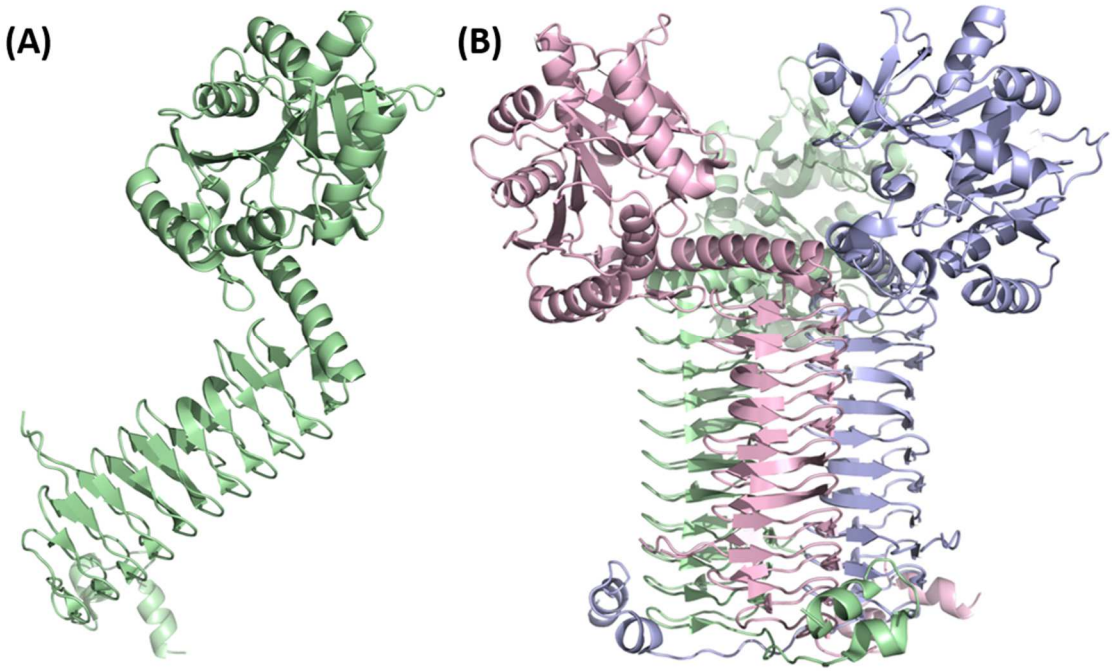


Figure 13. First reported structure of *M. tuberculosis* GlmU.

This figure depicts the *M. tuberculosis* GlmU (pdb 3DV8) quaternary structure. Each monomer is displayed as a cartoon representation of the secondary structure. (A) The GlmU monomer, showing a distinct two domain organisation on the polypeptide. The *N*-terminal domain is a nucleotide-binding Rossmann fold, while the *C*-terminal is a L*β*H fold. (B) A side view of the GlmU trimer, which shows the organisation of the three monomers around a threefold axis of symmetry. In addition, the extra-terminal region of each *C*-terminal monomer can be seen folding around the neighbouring partner in the trimer.

The *C*-terminal domain (residues 263-478) is comprised of L*β*H structure, with repeating isoleucine-rich hexapeptide motifs arranged as a triangular prism with a diameter of about 17 Å. The interior of the L*β*H is comprised of mainly hydrophobic residues, generally arranged in stacks of aliphatic or aromatic residues, which results in a long, narrow channel, aligned with the helix axis, measuring 2 - 3 Å. A cacodylate ion was observed on the threefold trimer axis, coordinated by Arg323 from each monomer. The Ac-CoA binding site is defined by the outer faces of two adjacent L*β*H domains, including an external, mobile loop that interrupts the eighth helical coil of one domain and a *C*-terminal tail donated by the third monomer of the trimer. The *M. tuberculosis* GlmU *C*-terminal domain has a *C*-terminal extension, comprising of two *α*-helices, which are linked by a series of *β*-turns. This *C*-terminal extension, which is 30-residues in length, differentiates *M. tuberculosis* GlmU from its orthologues, as it is significantly longer and is postulated to provide additional

stability to both the trimer and Ac-CoA binding site[215] (Figure 13 (B)). The acetyltransferase domain stabilises the polar groups of Ac-CoA with hydrogen bonds, in addition to making hydrophobic interactions with the large carbon chain of the cofactor. The adenine ring of Ac-CoA is bound by hydrogen bonds made with Ala434, as well as stabilisation by hydrophobic interactions by Ala451 and Val449. Further hydrogen bonds with Ser416, Trp460 and Lys464 are proposed to stabilise Ac-CoA, in addition to hydrophobic contacts with Gly390 and Val396. The thiol of Ac-CoA is stabilised by hydrogen bonds to Ala391 and Ser416. The aforementioned mobile loop contributes Tyr398 to the Ac-CoA binding site, which interacts with the acetyl and amino group of Ac-CoA. The proposed catalytic mechanism is like *S. pneumoniae*, *E. coli* and *H. influenzae* GlmU, in that the conserved active site His (His374, *M. tuberculosis* GlmU) deprotonates the amine group of GlcN-1P, followed by subsequent nucleophilic attack on the carbonyl carbon of Ac-CoA. The tetrahedral intermediate is then postulated to be stabilised by either an oxyanion Ala or Ser residue (Ala391 or Ser416, *M. tuberculosis* GlmU)[245].

The two domains are joined by a long *α*-helical arm of 22 residues, suggesting that GlmU evolved by fusion of an uridylyltransferase- and an acetyltransferase-encoding gene.

The *N*-terminal domain of *M. tuberculosis* GlmU shares a high degree of similarity with the uridylyltransferase domain of *S. pneumoniae* GlmU. The apoenzyme structure of *M. tuberculosis* GlmU reveals that the GlcNAc binding site of the uridylyltransferase domain is in a closed conformation. The holoenzyme structure with UDP-GlcNAc bound is very similar to the conformations observed in the *S. pneumoniae* GlmU structure. The alternate holoenzyme structure with GlcNAc-1P bound confirms that the interactions made with the GlcNAc moiety drive binding and the subsequent orientation in the uridylyltransferase active site. The later liganded *M. tuberculosis* GlmU structure led to the proposal that the order of substrate binding is UTP first, followed by GlcNAc-1P. The uridylyltransferase reaction mechanism is proposed to proceed after GlcNAc1-P binding, when the phosphate O atom initiates a nucleophilic attack on the *α*-phosphate of UTP, followed by a pentacoordinate intermediate at the *α*-phosphate of UTP stabilised by the Mg^{2+} ion. The final steps are believed to include the inversion of stereochemistry at the *α*-phosphate, UTP hydrolysis and release of PP_i. The uridylyltransferase active site Mg^{2+} is proposed to fulfil several roles in catalysis, including enforcement of the appropriate geometry for

catalysis through its octahedral coordination, polarisation of the P—O bond of the *α*-phosphate of UTP, which activates the phosphorus for nucleophilic attack and finally, to balance the negative charge of the transition state. There have been reports that the *N*-terminal domain of *M. tuberculosis* GlmU utilises two Mg^{2+} ions per active site, whereby Mg^{2+}_A enables nucleophilic activation and Mg^{2+}_B subsequently stabilises the transition state[246,247]. The Arg19 residue in the uridylyltransferase active site is important for the proposed catalytic mechanism as this residue is believed to bind the *α*-phosphate or the *β*-phosphate of UTP prior to the reaction. This Arg19 interaction would stabilise the phosphate charge and directing it away from the site to be occupied by GlcNAc-1P, leaving the *α*-phosphate open to direct attack from GlcNAc-1P[215].

1.3.5 *M. tuberculosis* GlmU essentiality

In order to identify conditionally essential genes in *mycobacteria*, Sassetti, *et al.* in 2001, developed a new technique, transposon site hybridisation (TraSH), which allows rapid functional characterisation by identifying the complete set of genes required for growth under different conditions[248]. Transposons are mobile DNA, genetic elements, that can change position within the genome and were first identified over seventy years ago by Barbara McClintock, who was awarded the Nobel Prize in 1983 for this discovery[249]. Transposition of these mobile genetic elements can create or reverse mutations, which alters the cell's genetic identity and genome size. Transposons have subsequently been used to develop techniques to identify essential and pathogenicity-related genes in microorganisms[250]. TraSH is a combination of transposon mutagenesis and microarray hybridisation, which is a technique that relies on hybridisation between two DNA strands. The probe DNA strands are immobilised on a small chip and subsequently hybridise target cDNA or cRNA. Probe-target hybridisation is detected and quantified by detection of either fluorophore-, silver- or chemiluminescence-labelled targets to determine relative abundance of nucleic acid sequences. The application of TraSH requires the use of a delivery system, such as a bacteriophage, in addition to a transposase gene, for instance Sassetti *et al.* used the C9 *Himar1* gene. Subsequently, large and diverse libraries of transposon insertion mutants are then generated in the target microorganism. Previous studies have shown that transposition is relatively random

aside from a requirement for the dinucleotide TA, which, for most mycobacteria genomes, would lead to saturating mutagenesis (Figure 14).

Transposon site hybridisation mutagenesis studies by Sassetti, *et al.* in 2003, to identify genes required for mycobacterial growth. This study identified genes that were subsequently assigned to essential pathways, where possible, as well as those that had unknown function. There were ten genes that were assigned to the *M. tuberculosis* peptidoglycan biosynthesis pathway and seven of these were identified as essential, including GlmU[251]. To confirm the TraSH study findings, analysis of the genes that were identified as essential were compared to the orthologous mycobacteria, *Mycobacterium leprae,* the causative agent of Leprosy. *M. leprae* has undergone reductive evolution, leading to a much-reduced genome consisting of orthologues of 40% of the functional genes found in the *M. tuberculosis* genome[252]. Most of the genes identified as being essential for *M. tuberculosis* were conserved in the degenerate genome of *M. leprae*, suggesting that non-essential functions have been selectively lost since divergence of the two species.

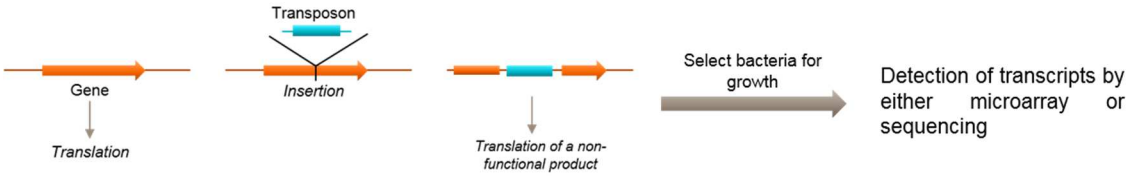


Figure 14. TraSH schematic.

The TraSH studies were followed by targeted, GlmU-specific studies carried out to investigate the essentiality of *glmU* for the growth of the *M. tuberculosis* orthologue, *Mycobacterium smegmatis,* which is often used as a surrogate for other mycobacteria, as it is both a "fast-grower" and non-pathogenic[253]. The *M. smegmatis glmU* gene was knocked out and growth was only observed in the presence of a temperature-dependent rescue plasmid coding for *M. tuberculosis glmU*. As the temperature of the growth conditions were altered, the transcription of *glmU* was reduced, leading to a reduction of the total amount of GlmU and a dramatic change in the morphology of *M. smegmatis*[244].

Recently, homologous recombination studies of *glmU* in *M. tuberculosis* demonstrated that both the acetyltransferase and uridylyltransferase activities of GlmU, through incorporation of K26A and H374A active site mutations, are

independently essential for bacterial survival *in vitro*. Additionally, GlmU depletion resulted in compromised survival of *M. tuberculosis* both *ex vivo* in a THP-1 cellular infection model, as well as *in vivo* in a guinea pig infection model[254]. During the THP-1 infection study, the impact of GlmU depletion was evident after 24 hrs and led to a significant reduction in survival of the pathogen 48 hrs post-infection. The *in vivo* guinea pig infection study highlighted that GlmU depletion prevented detection of *M. tuberculosis* upon doxycycline treatment, whereas discrete bacilli were observed in the control conditions. The pathology of the guinea pig lungs, upon termination of the study, displayed significant differences between the control and the GlmU depleted infections. Considerable granulomatous architecture was observed for the control infection and normal lung parenchyma observed for the *M. tuberculosis* GlmU mutant[254]. Scanning electron microscopy of *M. tuberculosis* bacilli depleted of GlmU in the study carried out by Soni *et al.* in 2015, displayed a similar deformed morphology to the GlmU mutated *M. smegmatis* bacilli observed by Zhang *et al.* in the aforementioned 2008 study[244,254].

In the past decade several independent, transposon insertion studies, using more sophisticated deep sequencing methodologies have been carried out to better characterise and understand the original TraSH study data presented by Sassetti, *et al.* in 2003. Deep sequencing is a next-generation sequencing (NGS) technique that is used to sequence a genomic region multiple times to detect small differences. These deep sequencing techniques, along with increased levels of transposon insertion saturation, have allowed greater understanding of sequence preference for transposon insertion. The application of the deep sequencing of transposon (Tn) insertion (TnSeq) approach, has allowed the assessment of the essentiality for *M. tuberculosis* growth of a comprehensive set of small genomic regions, including small ORFs, promoters, and small (noncoding) RNAs (sRNAs). These studies have all independently identified, and confirmed, that *glmU* is an essential gene for *M. tuberculosis* growth[255-257].

The results of the studies reviewed within this section, taken collectively, provide strong evidence that the *glmU* gene and both catalytic activities of the transcribed enzyme are essential for *M. tuberculosis* growth. Furthermore, to confirm these observations, site directed mutations in both the acetyltransferase and uridylyltransferase active sites, described in Section 1.3.4, have been shown to inactivate the respective catalytic activities of GlmU.

1.4 Anti-bacterial drug discovery

1.4.1 A history of antibiotic drug discovery

The discovery of novel pharmacologically active chemical substances has long been an arduous, protracted and invariably expensive interdisciplinary endeavour. Pharmacologically active chemical substances or novel molecular entities (NMEs), better known today as drugs, have been researched in earnest for over a century, ultimately culminating in the approval for use in USA of over 1453 (as of 2013) NMEs by the US Federal Food and Drug Administration (FDA)[258]. The years since 2013 have been particularly productive for NME approvals by the US FDA, with an average of 43 drugs approved per year over the five year period[259]. This increase in approvals has come at significant cost, which has been well documented[260], and is frequently reported as a decrease in productivity per billion US$ spent on pharmaceutical research and development (R&D)[261]. The decrease in pharmaceutical R&D productivity has come at a time when the approaches available to find novel chemical entities (NCEs) are, arguably more plentiful than ever before. In addition, the subsequent mechanisms to either inhibit or activate the drug target of interest are increasingly more innovative, for example the recruitment of the cellular protein degradation machinery to remove a protein of interest using bivalent small molecules known as proteolysis-targeting chimeras (PROTACs)[262].

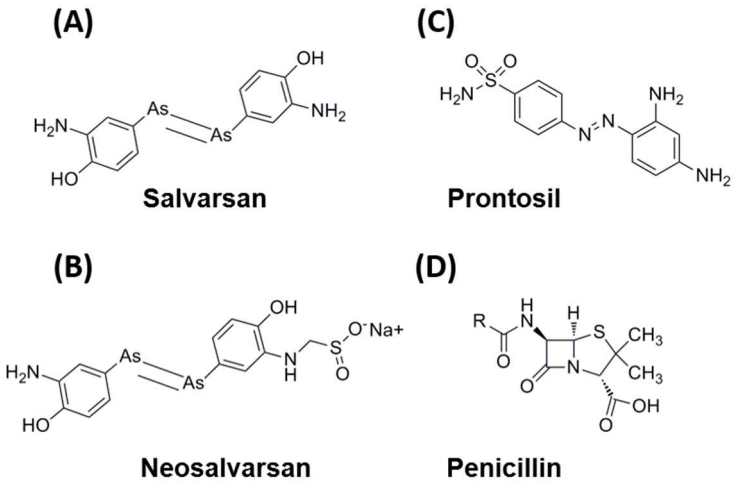


Figure 15. Structures of the early antibiotics.

The growth of several large pharmaceutical companies in the mid-twentieth century, was due to the need for industrial partners to help scale the production of penicillin,

one of the first antibiotics. The first antibiotics, originally known as “chemotherapy”, arsphenamine (Salvarsan) and neoarsphenamine (Neosalvarsan) were discovered by Paul Ehrlich, at the beginning of the twentieth century, to treat the bacterium *Treponema pallidum,* the causative agent of syphilis (Figure 15 (A-B)). These arsenic-containing antibiotics were the products of Ehrlich’s search for “magic bullets” that would attack the parasite while sparing the host[263]. Ehrlich was one of the first to describe the idea of receptors based on his interest in the immunology and chemotherapy of infectious diseases. His idea was that bacterial toxins combine with nutrient-capturing structures of cells, known as “sidechains”, which subsequently starved them. The cells respond by making more of these sidechains, some of which escape into the circulation as “antibodies” that combine with the toxin and make it harmless[264]. Paul Ehrlich was awarded the Nobel Prize for Physiology or Medicine in 1908, primarily for his work on both anti-diphtheria serum, as well as his discovery of Salvarsan[265]. Ehrlich’s discovery led to systematic screening of thousands of drug-like molecules at the German chemical company IG Farben (a legacy conglomerate that included Bayer AG), which led to the discovery of the antibiotic Prontosil (sulfamidochrysoïdine) by Gerhard Domagk, a German bacteriologist who was awarded the Nobel Prize in 1939[266] (Figure 15 (C)). The famous discovery of penicillin by Alexander Fleming and the subsequent work by Howard Florey and Ernst Chain, who developed the use of penicillin for medical applications, was recognised by the award of a joint Nobel Prize for Medicine in 1945[267] (Figure 15 (D)). The initial discovery of penicillin was in 1929 and by 1939 a team including Florey and Chain, were able to isolate enough penicillin to carry out a small clinical trial of the new antibiotic[268,269]. Due to World War II the British were unable to scale the production of penicillin and sought help from the USA and the US Department of Agriculture, where a team from the Northern Research Laboratory optimised the production of the compound and with the help of industrial partners, were able to produce enough of the antibiotic to satisfy the demands of the Allied Armed Forces within two years. The industrial pharmaceutical companies involved in the production of penicillin such as Merck (now Merck Sharp & Dohme or MSD), Pfizer, Abbott Laboratories and Squibb (now Bristol Myers Squibb) all continued to produce and sell the drug after World War II, starting what is known as the “Golden Era” of antibiotics, which is the period between the 1950s and 1970s. This period, further encouraged by Waksman’s discovery of Streptomycin, saw the exploitation

of natural product scaffolds and alternative versions of the aforementioned antibiotics, which were uncovered by mining the specialised metabolism of bacteria and fungi or by the chemical modification of existing scaffolds. This period of antibiotic-wealth, allowed researchers to shift their focus to other areas such as oncology, inflammation and neurosciences, while physicians, aided by antibiotics, were able to pioneer countless invasive surgeries, that today are viewed as routine procedures[270]. In addition to the discovery of novel antibiotics from microorganisms, several pharmaceutical companies utilised their microbiological capabilities to identify drugs that exerted other pharmacological or chemotherapeutic properties, including Lovastatin, a HMG-CoA reductase inhibitor and the immuno-suppressant Cyclosporin A, which was discovered in 1972 during an antimicrobial screening programme[271].

There has been a recent renaissance in natural product drug identification, which has culminated in the award of the 2015 Nobel Prize in Physiology or Medicine to William C. Campbell and Satoshi Omura, for the discovery of the microbial natural product avermectins, and to Youyou Tu, for the discovery of the plant natural product artemisinin. The derivative of avermectins, ivermectin, have lowered the incidence of the neglected tropical diseases onchocerciasis (also known as river blindness) and lymphatic filariasis (also known as elephantiasis). Artemisinin has significantly reduced the mortality rates for patients suffering from malaria and has an advantage over other malarial drugs, as it has a rapid rate of kill at all life cycle stages of the parasites[272]. In part, the advent of HTS and combinatorial compound synthesis has seen many pharmaceutical companies significantly scale back or abandon natural product screening. Additionally, issues with natural product compound collection purity, storage and supply issues have also played a part in the down-prioritisation of this drug discovery approach. However, recent advances in capabilities in the following disciplines; microbial genomics and metagenomics, metabolomics, synthetic biology, natural product biosynthesis, bioinformatics and analytical technologies, has prompted a growth in the utilisation of natural products with an increasing number of therapeutics in drug development, such as the anti-cancer chemotherapeutic, monomethyl auristatin E, a synthetic analogue of dolastatin 10 from the sea hare *Dolabella auricularia*[273].

The concept of target-based drug discovery is largely based on the inter-related disciplines of receptor pharmacology and enzymology. The discovery and

characterisation of the enzyme carbonic anhydrase was first described in 1933, was later followed, in 1949, by the discovery that sulphanilamide, the active metabolite of the sulphonamide, Prontosil, inhibited this enzyme[274,275]. This discovery lead to both better carbonic anhydrase inhibitors and the optimisation of the sulphanilamides as drugs, which are still used today to treat conditions as diverse as type II diabetes mellitus, edema and glaucoma[276].

Receptor pharmacology was originally based on the early work of John N. Langley and Paul Ehrlich and then later formalised by Alan J. Clark, who, in 1933, published a book entitled "The Mode of Action of Drugs on Cells", that quantified the relationship between drugs and receptors[277]. The quantification of these interactions were based on the mathematical approaches used in enzyme kinetics and based on the simple premise that, for many drug-receptor interactions, the relationship between drug concentration and biological effect followed a simple hyperbolic function[278]. A seminal manuscript by Raymond P. Ahlquist, in 1948, divided the adrenergic receptors, a class of G-protein coupled receptors (GPCRs), into the α- and β-adrenoreceptor subtypes[279]. The study presented in Ahlquist's paper thoroughly characterised the pharmacology of the adrenergic receptors in almost all organs, which helped characterise the activity of several existing drugs and also provided the basis of a large number of diverse future chemotherapeutics (over 117 approved agents by 2017), including β-blockers, β-agonists and benzodiazepines[280].

1.4.2 Drug discovery strategies

1.4.2.1 High-throughput diversity screening

Rapid advances in genomic sciences, widespread DNA sequencing, molecular biology, cell-based assays and combinatorial chemistry in the late-twentieth century led to the industrialisation of the drug discovery process by the pharmaceutical industry[281-285]. The combination of the advances in the aforementioned disciplines in most pharmaceutical and biotechnology companies, and also to a lesser extent in large universities and research institutes, was manifested in the application of high-throughput screening (HTS) to drug targets of interest. The HTS concept is essentially an "all-against-all" exercise, whereby a company or institute, will screen a drug target of interest, be it either an isolated protein, cell line over-expressing the target or a disease-specific primary cell, against their amassed collection of small

molecules. The assays used for HTS vary depending on the nature of the drug-target but in almost every case the intention is to apply the simplest, most robust and cost-effective methodology, regardless of detection technology used[286]. The assay is then used to screen a collection of compounds that can vary from thousands to millions and which is intended to cover as much of the hypothesised "drug like" chemical space as possible[287]. A HTS compound collection represents the chemical history of company or institute, based on their historical portfolio of internal projects, directed efforts to target small molecule libraries to specific areas of biological relevance, and synthesis or acquisition based on chemical diversity and synthetic tractability[288]. The success of the HTS strategy, set against the backdrop of decreasing pharmaceutical R&D productivity and increasing costs, has been extensively debated with no consensus reached[261,287].

An issue that has afflicted many HTS projects is an underlying false positive hit rate, caused by nuisance or pan-assay interference compounds (PAINS)[289]. These PAINS molecules can be artefacts; their activity does not always depend on a specific, drug-like interaction between molecule and protein. Nuisance compounds have non-ideal reactivities that appear to be drug-like binding, which yields false results across a broad cross section of assay technologies and drug targets. The observations of Baell, *et al.* have prompted a number of measures across the drug discovery community, including a set of guidelines from the American Chemical Society (ACS)[290]. At GSK, it has been necessary to employ a variety of filter strategies: sub-structures for reactivity, desirability, stability, and true PAINS, as well as molecular property, promiscuity (IFI), and sample provenance considerations[291]. In addition to nuisance filters, GSK has also prepared a set of a thousand known PAINS and other non-desirable molecules, termed the "nuisance set", which is screened prior to any hit identification screen to assess the susceptibility of both the molecular target and the assay to PAINS molecules.

1.4.2.2 Mechanism-based and knowledge-based drug discovery

Increasingly it is has become apparent that the complexities of the mechanism of action of the drug target of interest have been neglected in the reductionist HTS approach. Enzymes have progressively become an important drug target class, while enzyme inhibitors and inactivators now comprise roughly half of all marketed

drugs and have transformed human medicine[292,293]. The discovery of the HIV-1 protease inhibitors, saquinivir and indinavir, by rational design drug design, greatly contributed to the downgrading of HIV/AIDS from a fatal to a manageable disease[294]. In essence, enzymes are dynamic proteins and throughout their catalytic cycle they represent essentially different targets by consequence of the different substrate-, intermediate or product-bound states[295]. However, it is rare for an enzyme target to be thoroughly characterised, in terms of both the kinetic mechanism (the order of substrate addition and product release) and the chemical mechanism of catalysis, prior to initiation of inhibitor discovery projects[296]. In neglecting to thoroughly characterise the enzyme of interest, opportunities to rationally design substrate- or product-based inhibitors are missed, generally due to time pressures and over reliance on HTS or diversity compound screening. The failure to find progress-able small molecule inhibitors from HTS has led to a reinvigoration of interest in enzyme mechanistic studies, with many pharmaceutical and biotechnology companies increasing their focus on using more physiologically relevant enzyme kinetic studies during the lead identification and optimisation phases. Generating further information from enzyme inhibitor characterisation, such as measures of residence time, in addition to the traditionally determined half-maximal inhibitory concentration (IC_{50}) values, are of interest[297-299]. To fully maximise the benefit of full mechanistic characterisation of both the enzyme of interest and the subsequent inhibitors it is essential to utilise this information in combination with the complimentary approaches of biophysics and protein structural biology, to provide an integrated pharmacological profile[300]. This detailed information can then be utilised effectively to understand the translation of the pharmacological profile of the inhibitor into cell-based and phenotypic assays.

A natural progression of utilising increasing knowledge of the mechanism of drug targets, and more specifically enzymes, is rational or knowledge-based inhibitor design. This approach utilises knowledge of the specific chemical mechanism of the enzyme of interest, which allows design of mimetics of either substrate, transition-state or product inhibitors. Transition-state analogues, which are stabilised forms of the enzymatic transition state, bind tightly and have been extensively studied for a number of systems by Schramm, *et al.*[301]. The design of transition-state analogues requires extensive use of kinetic isotope effects to isolate the contributions of specific atoms in the substrate on the reaction rate, as well as both chemical and protein

crystallography. The complexity and slow timelines for developing transition-state analogue inhibitors has restricted this approach to academic studies, with few reports of pharmaceutical or biotech company activity. There are many more reports of inhibitory substrate analogues, from both industry and academia, with perhaps the most successful examples being both the HIV aspartic and Renin protease inhibitors[302,303].

1.4.2.3 Compound binding-based drug discovery

Another widely used method form of rational drug discovery is the use of structure-based inhibitor design, which can effectively be utilised in tandem with small molecule fragment-based screening (FBS)[304]. FBS involves the utilisation of sensitive biophysical techniques, such a surface plasmon resonance (SPR) or isothermal calorimetry (iTC), to detect binding of small molecule fragments to the protein of interest. Structural biology methods, such as X-ray crystallography or Nuclear Magnetic Resonance (NMR) spectroscopy, can then be used to rationalise compound binding and direct synthetic chemistry modifications to these small fragments[305]. This approach can rapidly and efficiently increase compound affinity for the protein target, while allowing other attributes of the small molecule, such as physio-chemical properties, to be effectively controlled from an early stage. This approach can be binding site agnostic, so increasing the chance of identifying novel and potentially allosteric binding hot spots on the protein target. An example of a successful fragment-based drug discovery project is a FBS for the *M. tuberculosis* pantothenate synthetase, which used a multi-biophysical assay screening cascade to initially screen in excess of a thousand fragments and triage these hits to identify molecules that bound to three distinct sites on the enzyme[306].

Another binding agnostic approach to drug discovery is a screening method known as DNA-encoded library technology (ELT), which is based on a concept first proposed in 1992[307]. This small molecule screening approach relies on the binding of compounds to the protein of interest rather than competition with a known binding partner or inhibition of an enzymatic activity. This allows the protein of interest to be probed with the DNA-encoded compounds in a non-biased manner, which offers the opportunity to identify novel compound binding sites. The ELT screening collection is comprised of small molecules that are unique as each molecule carries a DNA

sequence that specifically encodes its structure and serves as a template for its amplification[308]. The GSK ELT libraries, known as DNA-encoded libraries or DELs, were constructed using a combination of enzymatic and chemical synthesis in a split-and-pool format[309]. The DNA-encoded molecules are selected by binding to the protein target of interest, which is immobilised on a solid-state matrix prior to exposure to compounds. The compounds that bind to the protein, are then released by protein denaturation and identified by DNA sequencing. The compounds of interest identified from the DEL screening are then followed up by off-DNA chemical synthesis and protein binding affinity characterised using either an orthogonal biochemical or biophysical assay. At GSK, there are more than a hundred DELs, which comprise of billions of unique molecules that have been generated from over forty chemical reaction types. The optimisation of the ELT process has enabled multiple proteins to be screened in parallel, which allows, for instance, potential drug targets to be assessed for druggability. A recent study by Machutta, *et al.* used the GSK collection of DELs to assess proteins from *S. aureus*, *A. baumannii* and *M. tuberculosis*, to identify and prioritise the tractability of these potential drug targets[310].

1.4.2.4 Focussed compound set-based drug discovery

A contrasting strategy to the full diversity, HTS approach is to utilise sets of compounds that have been collated based on prior knowledge of either the drug target of interest, or related proteins. The contrasting approach is known as focussed screening and has been increasingly popular, due to improving computational chemistry techniques, as well as maturing knowledge of small molecule interactions with specific classes of drug targets, such as kinases, proteases and class A GPCRs[311]. Indeed, several of these focussed sets of small molecules have been made available to the wider scientific community by research institutes or companies, including the Structural Genomics Consortium (SGC)[312] and GSK. These sets are targeted against either epigenetic reader domains or kinases, with the later known as the "Published Kinase Inhibitor Set" (PKIS)[313]. As well as published focussed sets, several groups have reported screening strategies using proprietary knowledge-based collections[314]. The main advantage of using the focussed set screening strategy is that there are generally a much lower number of compounds than HTS collections, which makes this a much more resource- and cost-effective approach.

This can be particularly beneficial if the protein or cellular drug target is only available in small amounts, due to either difficulty in preparing the screening reagents, regulatory restrictions (such as higher biosafety levels) or cost. The use of increasingly complex cellular systems and more specifically either biologically relevant human or bacterial assays, make the use of focussed compound screening even more desirable.

As previously discussed, prior to the advent of the genomics era of the late twentieth century, drug discovery was primarily carried out using phenotypic assays. Recently, an analysis of FDA approvals between 1999 and 2008 demonstrated that phenotypic screening delivered more first-in-class drugs than molecular target-based approaches[315]. These data highlighted that phenotypic screening is gaining new momentum in drug discovery and this approach may help bridge the gap between R&D expenditure and productivity[316]. It should be noted that the same study highlighted that target based molecular screening strategies were more successful for follower drugs.

1.4.2.5 Phenotypic-based drug discovery

The definition of phenotypic screening is the use of a biological system, for example cells, tissue explant or whole organism, to identify small molecules that induce a desired, observable outcome, such as inhibition of cellular proliferation, cell death, induction of hormone secretion, amongst others. The benefit of utilising a phenotypic screening approach is that one can directly assess a test compounds ability to directly modulate a disease-relevant phenotype, for instance the up-regulation of an anti-inflammatory cytokine, such as interleukin 10 (IL10) from a human T cell, or activation of phagocytosis of bacteria by macrophages[317,318]. Phenotypic screening is a step beyond the binding site agnostic approaches of FBD and ELT screening, it is a black-box system, where the direct molecular target of the test compound is left as an unknown variable (or the test compound may modulate multiple proteins, which is known as network pharmacology or polypharmacology[319])[320]. The rationale for the current success of phenotypic screening is that the there is an unbiased identification of either the pharmacologically relevant pathway, the molecular target, or both leading to a more thorough understanding of the test compound mechanism of action[285]. In many cases, where novel inhibitors have been identified using a

phenotypic screening approach and subsequently the molecular mode of action (MMOA) has been determined, totally unprecedented mechanisms were characterised. A good example of this is the identification of a novel MMOA for reduction of PCSK9 in Chinese Hamster Ovary (CHO) cells, whereby the hit molecules were found to stall the ribosome in a gene-specific manner[321].

The challenge for phenotypic drug discovery is whether to proceed without any knowledge of the MMOA, or to deconvolute the screening assay to ascertain the mechanism. The risks of proceeding without knowing the MMOA are difficulties selecting clinical does, prediction of toxicity and understanding of on and off target toxicity. There have been many advances in molecular technologies that can be utilised to characterise the MMOA of novel inhibitors, such as chemoproteomics[322], genomics[323] and metabolomics[324]. An example of the deconvolution of a phenotypic screen to identify a novel MMOA, is from a screen to monitor overexpression of apolipoprotein A1 (ApoA1), a major component of high-density lipoprotein (HDL), from HepG2 cells, which identified novel small molecules[325]. The MMOA of these molecules were subsequently identified, using a chemoproteomic approach, as bromodomain with extra-terminal tail (BET) inhibitors[326].

However, despite the risks associated, it is not essential to know the MMOA of a phenotypically identified inhibitor while progressing through lead optimisation and into drug development. Perhaps the most notable example of a drug discovered and developed with knowledge of the MMOA is Aspirin, also known as acetylsalicylic acid, which is a medication used to treat pain, fever or inflammation[316]. Acetylsalicylic acid was discovered by chemists at Bayer in 1897 and subsequently marketed and sold as Aspirin from 1899. However, the MMOA was only discovered in 1971, by John Robert Vane (who was later awarded the Nobel Prize for his work[327]), who identified that Aspirin suppressed the release of prostaglandins and thromboxanes, through irreversible inhibition of the cyclooxygenase enzyme, COX-1. This identification of the Aspirin MMOA was 72 years after the first drug was sold to patients and the full mechanism is still not fully understood, as additional actions have been subsequently identified[328].

A disease translation challenge for phenotypic screening approaches is that the assay used must predict the clinical therapeutic response to a drug with a specific MMOA. The disease translation of phenotypic assays for the oncology and neuroscience therapeutic areas are difficult, mainly due to either the complexity and

heterogeneity of most late stage cancers, or the lack of predictive animal models for neuropsychological disorders. In direct contrast, phenotypic drug discovery approaches for infectious diseases where inhibition of the replication of bacteria, viruses or parasites in an assay, often have a strong translation to anti-infective activity in *in vivo* preclinical models.

1.4.2.6 Anti-infective phenotypic-based drug discovery

In recent years, there have been several reports of anti-infective HTS approaches, where large collections of small molecules are screened using phenotypic assays that monitor inhibition of growth of the test microorganism[329]. The nature of HTS screening collections, numbering between one to two million compounds, present several challenges, including the requirement of a high-throughput, microplate assay, in addition to counter-screens to control for non-specific cytotoxicity. From 2008 to 2010, three research organisations, Novartis[330], GSK[331,332] and St. Jude's Children's Research Hospital[333] published the results of *Plasmodium falciparum* phenotypic HTS drug discovery efforts, including compound structures, to provide novel starting points for anti-malarial researchers. GSK have been prolific in making HTS derived hit compound sets available to the wider scientific community[334], publishing the results of screens against *M. tuberculosis*[335,336] and three kinetoplastid parasites (*Leishmania donovani, Trypanosoma cruzi* and *Trypanosoma brucei*)[337].

Phenotypic drug discovery for anti-mycobacterial inhibitors have been successful recently, Bedaquiline (brand name Sirturo™, formerly TMC207), was first recognised as a potential drug in a whole-cell *M. tuberculosis* H37Rv assay. This potential anti-mycobacterial activity was confirmed against a number of *Mycobacterium* species, as well as drug-resistant strains of *M. tuberculosis*[338].

The GSK anti-mycobacterial phenotypic screen was carried out in multiple phases; first *Mycobacterium bovis* Bacillus Calmette-Guerin (BCG) was utilised to screen the 2 million compound GSK HTS collection at a single concentration to identify initial hits. The aim of the project was to identify a collection of chemical starting points for inhibition of *M. tuberculosis* growth. Prior to initiating the HTS, the GSK team screened a subset of the compound collection against both mycobacteria to validate the use of the *M. bovis* BCG assay as a surrogate for *M. tuberculosis* H37Rv. The data generated from the validation experiment were decidedly mixed, with only 55%

of the *M. bovis* BCG hits (identified screening the subset at 10 μM of test compound) able to inhibit *M. tuberculosis* growth. The validation experiment was also carried out at 25 μM of test compound and 86% of *M. bovis* BCG hits were confirmed as pan-active. These data highlighted that the planned screening strategy would identify a significant number of *M. bovis* BCG-only hits but the advantage of not having to carry out the HTS at biocontainment level 3, was deemed to be the more important factor. After the initial *M. bovis* BCG single concentration HTS and filtering of hit compounds through several chemical and physiochemical filters, a set of 3509 compounds were identified for further testing. The next phase of the HTS project was to, in parallel, determine IC_{50} values against *M. bovis* BCG and assess the selectivity of these hits by testing dose responses of compounds in a human HepG2 (an immortalised liver cancer cell line) counter-screen. The final stage of the hit discovery project was to test 777 *M. bovis* BCG, non-cytotoxic compounds in the *M. tuberculosis* assay to determine minimum inhibitory concentrations (MIC). The research team ultimately shared a set of 177 compounds, that could be broadly divided into seven chemical families, in Ballel, *et al* in 2013[335]. A further anti-mycobacterial HTS, utilising parallel testing of 254043 compounds in both *M. bovis* BCG and *M. tuberculosis* assays was carried out. This screen identified an additional set of 50 small molecules that were presented by Rebollo-Lopez *et al.* in 2015[336].

The set of 227 *M. tuberculosis* inhibitors have subsequently been used as a focussed small molecule set for target-based drug discovery to identify the molecular targets for these hits. Two recent reports have identified the molecular targets of the 227 *M. tuberculosis* "TB set" compounds; GSK3011724A was found to bind and inhibit β-ketoacyl synthase (KasA)[339], while a sulfolane and two indoline-5-sulfonamides were identified as tryptophan synthase inhibitors[340]. Additionally, an expanded collection of *M. tuberculosis* compounds, known as the "TB box" set, which consists of 10 000 small molecules (based on a selection of compounds from the original, GSK *M. bovis* BCG HTS) was screened in a modified *M. bovis* BCG assay, with inosine monophosphate dehydrogenase (IMPDH) over-expressed. This cell-based, target-specific screen identified seven inhibitors of the enzyme GuaB2 (Rv3411c), which is the only catalytically active and essential *M. tuberculosis* IMPDH (essential by TraSH)[341].

The ongoing innovations and information sharing in anti-infective drug discovery are urgently needed as TB, malaria, visceral leishmaniasis and Chagas disease, still

pose a threat to human lives, particularly in the developing world. The target product profiles set out for the discovery and development of new drugs for TB are particularly challenging. A new TB drug is required to shorten the duration of current treatment, demonstrate efficacy against drug-sensitive and drug resistant strains and show potential for use in drug combinations in developing countries[342].

1.4.3 Antibiotics and antibiotic resistance

An overview of antibiotics and the bacterial mechanisms of resistance have previously been discussed in Section 1.2.5, however the significance of this class of therapeutics and specifically, the discovery of new drugs is an important issue facing pharmaceutical R&D. Antibiotic resistance has been an issue from the first observations of the enzymatic degradation of penicillin[343], to recent studies that have shown the pathogenic bacteria, such as Enterobacteriaceae, which have become resistant to natural product drugs but also semi-synthetic penicillins, cephalosporins and newer carbapenems[344]. A review by Fischbach and Walsh, in 2009, classified antibiotic-resistant pathogens into three groups, all of which are major threats to public health. The first is methicillin-resistant *S. aureus* (MRSA), which is estimated to cause roughly 19 000 deaths per year in the United States. MRSA infections are estimated to cost to $3-4 billion per year and, in addition increases the likelihood of vancomycin-resistance *S. aureus* (VRSA), which is just as deadly but more challenging to treat. The second class was identified as multi-drug resistant (MDR) and pan-drug-resistant (PDR) Gram-negative bacteria. These pathogens that include strains of *A. baumannii, E. coli, K. pneumoniae, and P. aeruginosa* are less prevalent than MRSA, but pose threat of infections that are truly untreatable. MDR and PDR Gram-negative bacteria are resistant to some or all the commonly used antibiotics, including penicillins, cephalosporins, carbapenems, monobactams, quinolones, aminoglycosides, tetracyclines, and polymyxins. The final class of antibiotic-resistant bacteria comprises of MDR- and XDR-TB, which have been previously discussed as a rising health threat in the developing world[345].

The cost of antibiotic resistance is huge, most importantly for human health and life but also economically. There have been predictions that, in the next 30 years, 2.4 million people in the developed world will die from infections with resistant microorganisms, which will cost up to US$3.5 billion per year. Furthermore, many

low and middle-income countries already have high resistance rates, which are predicted to increase disproportionately[346]. A recent study found that burden of bacterial resistance infections is comparable to that of influenza, TB and HIV/AIDS combined and worryingly, the study found that 39% of all these resistant infections were caused by bacteria that could not be treated with even the last-line antibiotics[347]. There is a clear and pressing need for new antibiotics. There have been two historical lines of antibiotic discovery, which have previously been discussed; synthetic chemical efforts and isolation of new natural product scaffolds. A report from a group at GSK in 2007 detailed the relative failure of 70 HTS drug discovery projects, using a synthetic chemical collection, carried out on either individual targets, complete macromolecular biosynthetic pathways or whole-cells. The targets selected by GSK were identified using a genomic approach, after detailed evaluation of more than 300 genes[348]. This report, which focussed on the advent of the genomics-era in the late-twentieth century, provided a cautionary tale of an over-reliance on new advances. In addition, in the two decades following the revolution in target-based drug discovery, no new antibiotics have been discovered by this approach. A later manuscript from AstraZeneca, detailing their anti-microbial drug discovery efforts spanning the period between 2001 and 2010, which broadly confirmed the findings of the GSK team. This review highlighted that although it was possible to identify tractable chemical hits from target-based screens, the translation of these hit molecules to leads, capable of inhibiting *in vivo* bacterial growth, proved challenging[349].

The target validation of potential bacterial targets for drug discovery has relied heavily on genomic data, technology and innovation. The use of high-throughput techniques that can create precise deletions at the genome scale that have been used to identify essential genes in the following bacteria; *B. subtilis*[350], *E. coli*[351], *S. aureus*[352], *P. aeruginosa*[353] and *M. tuberculosis*[354]. The issue that has been identified with these gene essentiality studies, is that these are dependent on the growth conditions of the bacteria, as well as the genomic context can greatly influence the dispensability of the corresponding gene[355].

Strategies to identify new antibiotics have been proposed by several experts, with some similar common themes. There is broad agreement that there is a requirement for new drug discovery targets and novel chemical starting points, ideally based on unprecedented compound scaffolds. There are many risks in progressing new anti-

bacterial drug targets. The validation of novel targets must be thorough, ensuring that a range of growth conditions are considered, and there has been great progress in the fundamental biology research in academia, despite a continuing shortfall in funding. As well as thorough validation of novel drug targets *in vivo*, deep biochemical characterisation of the purified protein prior to any hit identification efforts. An understanding of the complex mechanisms of existing antibiotics, including the polypharmacology of compound classes such as the fluoroquinolones, could inform future drug discovery efforts and the design of new inhibitors. A recent study revealed previously unknown knowledge of the mechanism of an old antibiotic, D-cycloserine (DCS), first discovered in 1954. The antibiotic was believed to inhibit two sequential enzymes involved in the biosynthesis of peptidoglycan, Alr and Ddl[356,357]. However, in 2013 researchers at the Francis Crick Institute demonstrated that DCS is a time-dependent inhibitor of both *M. tuberculosis* and *E. coli* Ddl enzymes[358], while, in an additional study, DCS was recently shown to be phosphorylated during inhibition of Ddl. These studies demonstrate there is still much to learn about the mechanisms of even the most established antibiotics and that novel observations can offer unique insights that can be utilised to develop new, improved drugs[359].

A greater understanding of the permeability of bacterial cell walls may inform selection of compounds for screening based on the physical properties of known antibiotics and chemicals that overcome these barriers. Natural product scaffolds do not conform to the physio-chemical properties of conventional medicinal chemistry but are able to penetrate bacterial cell barriers and interact with antibiotic targets. There could be much to be learned from the better understanding of natural product antibiotics that may inform the choice of synthetic chemicals for screening. Finally, better prioritisation of potential antibiotic leads, prior to medicinal chemistry optimisation is required. The conventional assays to monitor a molecules ability to inhibit cell growth, leading to determination of an MIC, are simplistic and do not correspond to more complex features of bacteria, such virulence factors and aspects of metabolism[270].

1.4.4 The Tuberculosis drug discovery and development landscape

The WHO has called for a new global strategy for TB R&D, as the authors of the 2017 WHO report on Global Investments in TB R&D, state that complex research endeavours are required to end the TB epidemic[360]. An estimate by the Stop TB Partnership of an increase in investment of US$ 9 billion is required between 2016 and 2020. Indeed, the same partnership estimate a 5-year delay in TB R&D funding could lead to an additional 8.4 million TB cases and 1.4 million TB deaths by 2030, equating to incurred excess treatment costs of over US$ 5 billion. Shockingly, as one of the leading drivers of antibiotic resistance, not investing in TB R&D now, will ultimately lead to greater economic losses of in excess US$ 100 trillion by 2050.

The global TB R&D spend between 2005 and 2015 was US$ 6.3 billion, with an overall increase in annual funding from US$ 358 million to US$ 620 million. However, this expenditure on TB R&D was far from steady, with decreases in some years in the aforementioned period, which highlights the unpredictability in funding. Most of the TB R&D investment between 2005 and 2015 came from the public sector, with the largest funders being the US National Institutes of Health (NIH) and the Bill and Melinda Gates Foundation, who together contributed 37% of all money spent on global TB research. Annual pharmaceutical industry investment in TB R&D peaked at US$ 145 million in 2011 before falling to US$ 87 million in 2015. This decline in TB R&D investment from the pharmaceutical industry has largely been due to the withdrawal of major multinational companies such as Novartis, AstraZeneca and Pfizer from active R&D between 2012-2014. Encouragingly, there has been a recent increase in TB R&D investment with funding exceeding US$ 700 million in both 2016 and 2017[361]. However, this increase in investment is still only 35% of the annual expenditure required to achieve the total investment of US$ 9 billion by 2025. The current levels of investment, impacted by the decline of industry investment and flat funding from public sector investors highlight the need for new resources in the TB research field and the need to develop innovative, flexible and collaborative mechanisms for advancing the science needed to end TB. There are learnings from history that give hope in the fight against TB, specifically in the second half of the twentieth century where the power of research to better public health was at its most productive. TB research advanced rapidly progressed throughout the 1940s and 1950s, taking TB treatment from sanatoriums, which were medical facilities for long-

term illnesses, into an antibiotic age[362]. As previously discussed in Section 1.1.6, the introduction of multiple new antibiotics from the 1950s to the 1970s culminated in the short-course DOTS combination chemotherapy that is still in use today. The advances in anti-tubercular drug therapies were complemented by earlier technologies, including the BCG vaccine[363], which was introduced in 1921, and diagnosis of TB through acid-fast sputum smear microscopy, developed in the late nineteenth century[364]. TB research has contributed to biomedical science in many ways, for instance the development of the BCG vaccine was one of the first demonstrations of a powerful new technique for vaccine creation, known as attenuation via serial passage[365]. The Medical Research Council (MRC) in the United Kingdom was originally founded in 1901 to investigate the links between human and animal TB[366]. In 1948, in another TB research contribution to medical sciences, the MRC TB unit organised and carried out the world's first randomised and controlled drug trial, investigating streptomycin to treat TB[367].

As outlined in Section 1.1.6, the majority of the drugs that are still used today as first and second-line drugs were discovered in the 20 years after the approval of streptomycin to treat TB[368]. The MRC TB unit in collaboration with the US Public Health Service (USPHC) developed the current successful but flawed treatment regimens that rely on drug combinations to decrease and control the emergence of resistant strains of *M. tuberculosis*[47]. Counter-intuitively, TB research suffered for its early success, with investments in R&D declining rapidly in the 1970s and 1980s, culminating in the MRC and USPHS disbanding their TB research programs in 1986[369]. The funding for TB R&D declined almost to nothing in the late-1980s and early-1990s with only modest annual NIH investments of US$ 1-2 million to individual researchers[370]. The situation did not improve significantly until 2000, based on reports of the global increase in incidences of TB cases and deaths (the spread of HIV, especially in Africa and the emergence of MDR-TB)[371], when the UN published the millennium health goals and the creation of a global fund to fight HIV/AIDS, Tb and Malaria[11].

Today, TB drug discovery and development is more successful than at any point in the previous 50 years with a number of new molecules and therapeutic combinations in preclinical and clinical trials[372]. The first novel treatments since the release of Ofloxacin in 1980, delamanid and bedaquiline, were approved for use as treatments for drug resistant TB in 2012 and 2013[373,374]. These two drugs have been fast

tracked with limited efficacy data and known serious safety issues, such as instances of QT prolongation that can lead to abnormal and potentially fatal heart rhythm[65].
The criteria for developing new TB drug candidates comprises a fully validated safety profile, greater potency than existing treatments, inhibition against a new target, as well as a compatibility with antiretroviral therapy (ART) and established or new TB drugs. The aim is to shorten the duration of TB treatment, demonstrate effectiveness against resistant strains of *M. tuberculosis*, enable co-treatment with patients also infected with HIV and provide the opportunity for future novel drug regimens[375]. Many of the drug candidates in clinical trials are repurposed molecules that are used to treat other infectious diseases, such as fluoroquinolones, which are inhibitors of DNA gyrase and DNA topoisomerase in many bacteria[372]. The other drugs in pre-clinical and clinical trials are new, TB-specific small molecules, many of which were discovered as anti-mycobacterial agents and have novel mechanisms of action, such as diarylquinoline (bedaquiline) that inhibits the *c* subunit of ATP synthase[376].
Target-based TB drug discovery can be successful if the drug target of choice has been shown to be essential for *M. tuberculosis* growth, such as the enzyme decaprenylphosphoryl-D-ribose oxidase (DprE1)[377] and the transporter Mycobacterial membrane protein Large 3 (MmpL3), which mediates the transport of substrates across the cell membrane[378]. The aforementioned targets have been extensively studied and there are a number of diverse small molecule series that have been identified as potential drugs, such as the benzothiazinones and diamine derivatives[379]. This precedent offers the opportunity to screen other essential *M. tuberculosis* proteins against targeted compound sets, in order to identify new chemical starting points for these potential drug targets and then directly optimise these molecules using target-based approaches-[380].

1.4.5 GlmU drug discovery

The identification of attractive targets for antibacterial drug discovery requires consideration of several criteria, which, while not guaranteeing success, may mitigate any progression issues with the lead molecules selected. The main target-based considerations are essentiality (at least *in vitro*), lack of close human homologues and absence of target-based cross-resistance. GlmU is an attractive antibiotic target, primarily due to its importance in bacterial cell wall biosynthesis, in

addition to studies that have shown its essentiality in both *M. tuberculosis* and *E. coli*, as well as the absence of any comparable enzymes in humans. There have been several reports that have shown the success in targeting both domains of GlmU from different bacteria, with small molecule inhibitors. A novel *H. influenzae* isoenzyme inhibitor, with a relatively modest IC_{50} of 18 μM, was identified by a HTS of the Pfizer compound collection targeting the uridylyltransferase domain and was later found to bind, using X-ray crystallography, a previously undescribed allosteric site on the *N*-terminal domain(pdb 2W0V and 2W0W)[214,225].

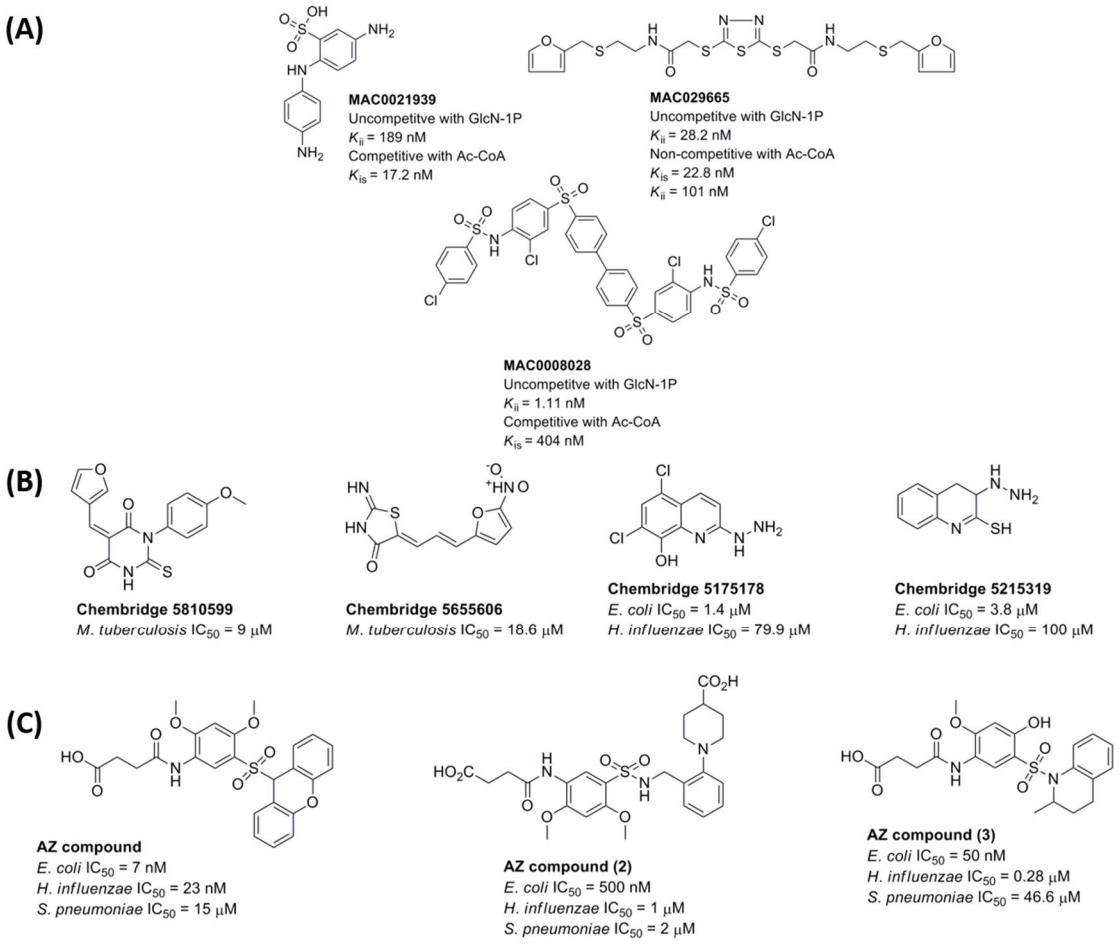


Figure 16. Exemplars of reported GlmU acetyltransferase inhibitor series.

Small molecule inhibitors of GlmU acetyltransferase activity of various isozymes. (A) Three lead compounds from a screen carried out by Pereira, *et al.* against *E. coli* GlmU[381]. (B) GlmU acetyltransferase inhibitors identified by screening a commercially available compound library (Chembridge) in two separate studies by researchers at the CSIR-Indian Institute of Integrative Medicine[382,383]. (C) GlmU acetyltransferase inhibitors identified by researchers at AZ and reported in three separate manuscripts; (1)[221], (2)[384] and (3)[385].

Inhibitors of *E. coli* GlmU acetyltransferase were identified from a HTS of 50 000 small molecules, using a coupled, assay that monitored the release of PP_i after uridylyltransferase activity. This HTS identified an initial set of 63 hits, which upon further deconvolution confirmed that 37 of these compounds were acetyltransferase activity inhibitors. Subsequent confirmation screening and mode-of-inhibition studies identified that the hit compounds were competitive with Ac-CoA and dependent on the presence of GlcN-1P. These findings, for the *E. coli* isoenzyme, suggest that the acetyltransferase kinetic mechanism progresses by initial binding of GlcN-1P, followed by Ac-CoA, which is contrary to structural biology studies that proposed that the mechanism follows the reverse order of substrate binding[220]. Kinetic mechanism studies determined for the other GlmU orthologues confirm that the acetyltransferase mechanism is ordered bi-bi, but the order of substrate binding agrees with the later observation, which is Ac-CoA binding first followed by GlcN-1P. The authors of the *E. coli* GlmU acetyltransferase HTS hit identification study stated that none of the three lead compounds were able to inhibit growth of *E. coli* MC1061 in culture[381] (Figure 16).

There have been further HTS studies, carried out by researchers at AstraZeneca, probing the acetyltransferase of *E. coli* GlmU, which identified a series of inhibitors, containing a sulphonamide group (Figure 16). The binding of these sulphonamide-containing inhibitors to the Ac-CoA binding site were confirmed using several biophysical techniques, including NMR and isothermal titration calorimetry. Optimisation of the sulphonamide-containing inhibitors lead to a 300-fold increase in potency of this chemical series against another Gram-negative GlmU, from *H. influenzae* but remained weaker or inactive against Gram-positive isozymes, from *S. pneumoniae* and *S. aureus*. The sulphonamide-containing inhibitors were shown to be anti-microbial in a strain of *H. influenzae*, which was lacking its major efflux pump. The inhibition of UDP-GlcNAc incorporation into bacterial macromolecules was shown to be consistent with anti-microbial activity being caused by disruption of peptidoglycan and fatty acid biosynthesis. However, anti-microbial activity against wild type strains of *H. influenzae* and a strain of *E. coli* lacking its AcrB-TolC efflux system suggested that the compounds were poorly cell permeably[221].

In parallel, the research team at AstraZeneca, carried out a HTS against the uridylyltransferase activity of *H. influenzae*, which identified an aminoquinazoline inhibitor series. This series of compounds were confirmed to be GlmU

uridylyltransferase inhibitors using secondary assays and medicinal chemistry optimisation led to several exemplar inhibitors that were equipotent against both *H. influenzae* and *E. coli* isozymes. Structural biology studies using with *H. influenzae* GlmU and aminoquinazoline inhibitors identified that this series of compounds occupy the UTP binding pocket in the uridylyltransferase active site. These X-ray crystallography experiments, in combination with enzymology and biophysical studies confirmed that the aminoquinazoline inhibitors are UTP competitive. Anti-microbial studies to determine MIC values using aminoquinazoline compounds were not successful using either the wild type *H. influenzae* and *E. coli* or the ArcB or TolC efflux pump mutants of both species[386]. Subsequently the aminoquinazoline inhibitor series were optimised through a combination of medicinal chemistry techniques, including virtual screening, increased potency of these compounds 1000-fold but failed to confer any significant anti-bacterial activity[385].

As previously mentioned, Payne *et al.* published a review in 2007 of anti-bacterial early drug discovery projects at GSK between 1995 and 2001. In this period 67 HTS projects were carried out using the legacy SmithKline Beecham compound collection (which consisted of 260,000 – 530,000 compounds) against several *S. pneumoniae*, *S. aureus* and *H. influenzae* targets, which had previously been identified as essential *in vitro*. As part of these extensive anti-bacterial drug discovery efforts, both activities of *S. aureus* GlmU were explored through three separate HTS studies, which included two acetyltransferase screens. The stated outcome of these HTS projects was that no progress able hit compounds of either of the activities of *S. aureus* GlmU were identified[348].

There have been few *M. tuberculosis* GlmU inhibitor screening efforts and one study identified hit compounds that inhibited acetyltransferase activity, with IC_{50} values ranging from 9 to 70 μM. Anti-bacterial activity of two lead compounds were determined in both drug-susceptible and -resistant *M. tuberculosis* strains, as well as showing increased MIC values in combination with rifampicin, isoniazid and ethambutol. Further studies identified that one of the lead compounds was cytotoxic in a eukaryotic cell line, while the other had previously been reported as an inhibitor of a number of other *M. tuberculosis* enzymes[382].

Structural alignments of the GlmU analogues revealed that the *N*-terminal domain allosteric site identified from the *H. influenzae* isozyme was also present on the *M. tuberculosis* protein. Virtual screening identified a lead compound that was

subsequently synthesised and shown to inhibit GlmU uridylyltransferase activity, as well as stabilise the *M. tuberculosis* enzyme T_m in DSF biophysical studies. The *M. tuberculosis* GlmU inhibitor was then shown to inhibit bacterial growth of the H37RV strain and to a lesser extent an engineered bacterial cell line that over-expressed GlmU. However, the allosteric *M. tuberculosis* GlmU inhibitor was shown to be poorly efficacious and the study failed to show any additional related inhibitor studies to support the observations made with the lead molecule[254].

The *M. tuberculosis* GlmU protein prepared for this PhD project was included in a large-scale ELT binding screen project against 161 proteins from *A. baumannii*, *S. aureus* and *M. tuberculosis*. This study ranked GlmU as a low tractable drug target, which was due to the low number of specific binding compounds identified. Interestingly, a related L*β*H acyltransferase, LpxD from *A. baumannii* was ranked as a tractable drug target and a novel chemical series was identified, which had *in vitro* anti-bacterial inhibitor activity[310]. The *M. tuberculosis* GlmU ELT binding compounds were followed up in an acetyltransferase activity assay and the results of this experiment are presented in Chapter 5.

Collectively, the studies reviewed in this section, demonstrate that GlmU is an attractive, broad-spectrum anti-bacterial drug target. The identification of a selection of novel GlmU small molecule inhibitors, against both acetyltransferase and uridylyltransferase activities, demonstrates that both active sites are moderately druggable and once compound series are identified the potency can be optimised through a range of medicinal chemistry approaches. The essentiality of both GlmU enzyme activities in *M. tuberculosis, E. coli* and *H. influenzae* were confirmed by novel, specific GlmU inhibitors. Inhibition studies using the aforementioned inhibitors of either the GlmU acetyltransferase or uridylyltransferase functions led to *in vitro* anti-bacterial activity. There are many other criteria that need to be satisfied before GlmU is fully validated as a good anti-bacterial drug target. What remains to be fully validated is the selectivity and safety of GlmU inhibitors in higher mammals and humans. Additionally, the susceptibility of GlmU to target-based resistance and subsequent rapid resistance selection has yet to be characterised. However, the evidence presented in this section demonstrates that both the activities of GlmU are essential, so targeting both functions may be a strategy to prevent antibiotic resistance from active site mutations. In addition, the presence of one potential allosteric binding site situated on the *N*-terminal domain of GlmU may provide an

opportunity to identify and optimise small molecules that elicit an inhibitory effect that do not directly bind to the uridylyltransferase active site. This review of GlmU inhibitors highlights that there are still chemistry challenges, which are universally applicable to all antibiotics, including permeability of bacterial cells to small molecules and, once these compounds are in the cell, avoiding bacterial efflux pumps[387-389].

1.5 Project aims and objectives

The central hypothesis of this project is that the acetyltransferase activity of GlmU is essential for optimal *M. tuberculosis* growth. The mechanism of GlmU acetyltransferase activity has not been studied kinetically and there are many unknown or unproven aspects, including, amongst others, the order of substrate of binding and product release, and the identity of the key residues for binding and catalysis. The initial aim of this project is to utilise, where possible, continuous assay approaches to characterise both the kinetic and chemical mechanism of GlmU catalysed acetyl transfer. The further aim of this project is to utilise the characterisation of the acetyltransferase activity to develop sensitive microplate based, kinetic screening assays to identify GlmU inhibitors. The final aim of this project is to utilise orthogonal approaches to confirm the GlmU hit compounds, prior to assessment of these compounds as *M. tuberculosis* inhibitors. These main project aims can be divided into three distinct segments of study that are detailed in the following sections.

1.5.1 Characterisation of the kinetic mechanism of GlmU catalysed acetyl transfer

The initial aim is to utilise the expression and purification methods previously reported for GlmU and orthologues to produce soluble, full length protein. Purified protein will then be quantified, prior to configuration of kinetic assays for both acetyltransferase and uridylyltransferase activities and initial determination of steady-state parameters for both functions of GlmU. The role, specificity and requirement for divalent metal ions during GlmU catalysed acetyl transfer will be investigated. Initial velocity pattern studies will then be carried out to ascertain the likely kinetic mechanism of GlmU acetyltransferase activity. The Ac-CoA and GlcN-

1P substrate specificity will be assessed to understand the tolerances of the acetyltransferase active site for close analogues. The acetyltransferase products will be assessed to identify whether, at high concentration there is any inhibition of catalysis. Product and dead-end substrate analogues will be utilised for inhibitor pattern studies to ascertain the likely order of substrate binding and product release. Finally, structural biology experiment using X-ray crystallography will be utilised to compliment the kinetic studies and help elucidate the critical interactions made in a dead-end mimic of the ternary complex.

1.5.2 Characterisation of the chemical mechanism of GlmU catalysed acetyl transfer

The effect of varying pH on GlmU acetyltransferase Michealis-Menten kinetic parameters, using a kinetic assay, will be studied to assess the role of general acid-base chemistry in catalysis and substrate recognition. To identify the ionisable groups involved in GlmU acetyltransferase activity, 1D NMR studies will be carried out to determine the pK_a value of the GlcN-1P amine group under the standard kinetic absorbance assay conditions, while varying the pH conditions. Site-directed mutagenesis studies will be carried out to investigate the effect of mutating important acetyltransferase active site residues. The mutant GlmU enzymes will be assessed by carrying out pH-rate studies to ascertain the effect on Michealis-Menten kinetic parameters, when compared to the WT (WT) profiles. will allow identification of the expression and purification methods previously reported for GlmU and orthologues to produce soluble, full length protein. Finally, Solvent kinetic isotope effect studies will be carried out to investigate the acetyl transfer chemical mechanism by identifying whether ^{1}H transfer is involved in catalysis, in addition to ascertaining whether either the solvent, or exchangeable groups on either GlmU or the substrates act as a ^{1}H donor.

1.5.3 GlmU acetyltransferase small molecule inhibitor identification and hit qualification

The characterisation of GlmU acetyltransferase activity will provide parameters for the development of robust, sensitive microplate-based screening assays that will be validated using GSK training small molecule sets. Subsequently, these GlmU

acetyltransferase assays intent will be used to screen small molecules form several different compound collections, both diverse collections of non-targeted compounds, in addition to focussed sets that have previously been shown to inhibit the growth of *M. tuberculosis*. Any hit compounds identified will be initially confirmed using titrations of test small molecules to allow quantification and ranking of inhibition by determination of IC_{50}. To further confirm inhibiting compounds, orthogonal biophysical assays will be utilised to either monitor both enzymatic activities in a label-free system or determine direct binding to GlmU. The direct label-free detection of both enzymatic activities will be enabled by use of the Rapidfire high throughput mass spectrometry methodology, while techniques such as Differential Scanning Fluorimetry (DSF) and MicroScale Thermophoresis (MST) will be utilised for direct binding studies. Finally, whole cell *M. tuberculosis* assays will be used to determine MIC values for novel GlmU inhibitors identified in the previous screening studies.

Chapter 2. Materials & Methods

2.1 Materials

2.1.1 General materials

All chemicals were of analytical or reagent grade and were used without further purification unless otherwise stated. All solutions were prepared using water from an 18.2 MΩ·cm Milli-Q synthesis A10 Q-Gard system which was filtered through a 0.22 µm filter. CoA-SH thioesters, Glucosamine 1-phosphate, buffers, and all other chemicals were of analytical or reagent grade and were purchased from either Sigma-Aldrich or Fisher Scientific, unless otherwise stated. All chromatographic columns were obtained from GE Healthcare.

2.1.2 UDP-Glucosamine preparation

UDP-Glucosamine (UDP-GlcN) was prepared by Martin Rejzek and Rob Field from the Molecules from Nature laboratory, John Innes Centre, Norwich. The chemical synthesis of UDP-GlcN was performed as published by Morais *et al.*[390], with some modifications, primarily, a change in the phosphorylation step was utilised. This change replaced the use of butyllithium to deprotonate the hemiacetal 1, followed by dibenzyl chlorophosphate, with the application of lithium diisopropylamide (LDA) followed by the use of the commercially available tetrabenzyl pyrophosphate, to generate dibenzylphosphate 2. The yield of UDP-GlcN was 82.0 mg, which was obtained in the form of diammonium salt after a single-step purification in an overall yield of 8.5% over seven synthetic steps. The analytical data were in good agreement with the literature precedent[390].

2.1.3 Knowledge-based inhibitor design and preparation

Novel GlcN-1P analogues, modified at the C-2 amine position, based on the proposed transition state, or simplified versions of the tetrahedral intermediate, were designed. Two GlcN-1P analogues, were designed to conserve some of the interactions made by CoA-SH, GSK217A (sulfonamide) or GSK3536777A (truncated CoA-SH sulfonamide). Three GlcNAc-1P analogues were designed to investigate the consequence of varying amide substitution length and steric bulk; GSK436A,

GSK017A and GSK930A. The final GlcNAc-1P analogue prepared was designed to contain a truncate of the carbon backbone believed to be present in the tetrahedral intermediate, GSK439A. The chemical structures of these knowledge-based inhibitors can be found in Figure 119. The design and synthesis of these novel GlcN-1P analogues was carried out by Ben Whitehurst, a PhD candidate in the Department of Pure and Applied Chemistry at the University of Strathclyde. The full synthetic routes for the preparation of these molecules is detailed in Ben Whitehurst's PhD thesis, entitled, "Development of Novel Inhibitors of Carbohydrate-Processing Targets Involved in *Mycobacterium tuberculosis* Cell Wall Biosynthesis".

2.2 Methods

2.2.1 General methods and equipment

Sonication was used to lyse *E. coli* cells, prior to ultracentrifugation, during protein purification, utilising a Sonics VCX750 Ultrasonic processor (Sonics & Materials Inc.) equipped with a medium high gain probe. Cell pellet was resuspended by stirring in a lysis buffer, then, while keeping the cell slurry on ice, sonicating for a total of 10 minutes, using cycles of 9 second pulses at 60 % amplitude, followed by 9 second cooling period.

Ultracentrifugation was utilised, prior to chromatography steps during the protein purification process, to remove unbroken cells, lipids, and particulate matter from the lysate solution. Lysate was equally divided into 40 ml centrifuge tubes, which were accurately balanced and added to an Avanti J-30I Ultracentrifuge (Beckman Coulter), in a JA-30.50 rotor, before centrifugation at 100000 x *g* for 30 mins while constantly maintaining the temperature at 4 °C. Subsequently, the clarified lysate was pooled on ice, taking care not to disrupt the pelleted cell material.

All chromatography steps during the protein purification process were carried out at 4 °C using an ÄKTA purifier 100 (GE Healthcare) equipped with a FRAC 950 fraction collector.

Chromatography fractions were assessed by sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel electrophoresis using a method derived from the original Laemmli protocol were applied.[53] Fractions, selected from chromatography traces where an increase in the absorbance at 280 nm was observed, were added to a PCR plate and diluted with SDS-PAGE sample buffer [62.5 mM

Tris(hydroxymethyl)aminomethane (Tris-HCl) (pH 6.8), 2 % SDS, 25 % glycerol, 100 mM Dithiothreitol (DTT), 0.1 % bromophenol blue], before incubation at 95 °C using a heat block for 10 mins to denature the samples.

Denatured samples were added to 4-20% Tris-Glycine Mini SDS-PAGE gels (Novex), submerged in Laemmli buffer and exposed to a constant voltage of 125 V, at 50 mA for 40 mins using a XCell SureLock Mini-Cell System (Thermo Fisher). In order to visualise protein molecular weight ranges during electrophoresis, SeeBlue® Plus2 Pre-Stained (Thermo Fisher) was added to the outermost wells of all gels. Following electrophoresis, SDS-PAGE gels were stained, while gently shaking, for 10 mins with Instantblue Coomassie reagent, before washing in H_2O and gel image capture using a G:Box Chemi XT Analyser (Syngene).

Non-reducing and native gel electrophoresis were used to analyse purified protein products, to identify either the presence of disulphide linkages or higher order oligomeric states, respectively. The non-reducing gel electrophoresis protocol is very similar to the reducing gel method, with the notable omission of 100 mM DTT from the SDS-PAGE sample buffer. The native gel electrophoresis protocol involved 5 µg of protein being prepared in Tris-Glycine native sample buffer [100 mM Tris-HCl (pH 8.6), 10 % glycerol, 0.0025 % bromophenol blue] before loading on a 4-20% Novex native PAGE gel (Thermo Fisher Scientific, cat. no. XP04200BOX) alongside Native Mark standards (Life Technologies LC0725). The gel was loaded in a XCell SureLock Mini-Cell System, submerged in Tris-Glycine native running buffer [25 mM Trizma (Tris-base) (pH 8.3), 192 mM Glycine] and exposed to a constant voltage of 125 V, at 50 mA for 2.5 hrs, at 4 °C.

Protein concentrations of homogeneous preparations were measured using the Bradford colorimetric protein assay[391]. Pooled SDS-PAGE fractions were titrated 1 in 2 from stock, in duplicate, leaving a final volume of 10 µl in a clear 96-well microplate. A standard titration of bovine serum albumin (BSA) from 1 mg/ml to 0.2 mg/ml was added in 10 µl volumes, in duplicate, to the same microplate, as well as 10 µl duplicates of the protein storage buffer, followed by additions of 250 µl Pierce Coomassie reagents to all sample containing wells on the test microplate. The 96-well test microplate was then transferred to a SPECTRAmax Plus (Molecular Devices) and absorbance measured at 595 nm. The absorbance at 595 nm of the standard curve was fitted to a four-parameter quadratic curve fit and these data are then used to calculate the protein concentration of the test samples. In the case that

the pooled SDS-PAGE protein fractions concentration is in significantly in excess of 1 mg/ml, the test sample is then diluted 1 in 10 in storage buffer, prior to titration and measurement of absorbance at 595 nm.

Protein, when necessary, was concentrated in Amicon Ultra filters with Ultracel-10 membranes (Merck Millipore) by centrifugation at 4000 rpm in an Eppendorf 5810R centrifuge, while maintaining the temperature at 4 °C. Samples were initially centrifuged for 10 mins, before mixing by pipetting to ensure that protein doesn't stick to the membrane. The centrifugation and mixing cycle is then repeated until the sample volume is reduced to a level corresponding to the desired concentration.

Purified GlmU was analysed by electrospray ionization mass spectrometry (ESI-MS) for accurate determination of the monomer mass and peptide mass-fingerprinting (PMF) to confirm protein identity.

All spectrophotometric and fluorescence assays were monitored at 30 °C using Tecan Safire 2, Tecan M1000 Infinite or Tecan M1000 Infinite Pro quad4 monochromator-equipped microplate readers. Prior to measurement of assay signals on the aforementioned plate readers, microplates were spun at 1000 rpm for 1 min using an Eppendorf 5810R. Absorbance assays were carried out exclusively in clear-bottom, black polystyrene, 384-well microplates (Greiner Bio-One cat. no. 781096), while fluorescence assays were performed in small volume, black polystyrene, 384-well microplates (Greiner Bio-One cat. no. 784076).

Differential Scanning Fluorimetry (DSF) experiments were carried out using a Roche LightCycler 480 II real-time polymerase chain reaction (RT-PCR) instrument equipped with a 384-well thermal block and performed in sealed Accuflow Roche 480, white, PCR, 384-well microplates.

All concentrations of GlmU WT and SDM mutants reported are the final concentrations used based on the molecular weight of the monomer.

To add reagents to microplates two liquid handlers were used, either the Multidrop Micro (Thermo Fisher), which was utilised for bulk reagent additions while screening small molecules, or the dragonfly discovery (TTP LabTech) that was employed for titrations of aqueous reagents and detailed mechanistic studies.

2.2.2 GlmU wild type plasmid preparation

The Rv1018c gene sequence from *M. tuberculosis* H37Rv was codon adapted for optimal protein expression in *E. coli*, and its nucleotide sequence was synthetically prepared and ligated into the *pJ411* plasmid (DNA 2.0). The GlmU DNA sequence was confirmed upon preparation and delivery to GSK by Sanger sequencing. Two GlmU constructs were designed and prepared; one containing a N-terminal hexahistidine affinity purification tag separated from the protein sequence by a Tobacco Etch Virus (TEV) protease sequence and another construct with a C-terminal hexahistidine affinity purification tag preceded by a TEV protease sequence. In order to confirm the DNA sequence, Sanger sequencing was used, utilising the primers, which were produced by Sigma-Aldrich and prepared as 100 µM stocks in RNAse free H_2O, listed below:

GlmU 441 forward - ATTGCGACCCACCGTGCGG
GlmU 901 forward – TCGGCCGTGACACCGTGATT
GlmU 1320 forward – TTTGTTGCGCCGGTCACG
GlmU 441 reverse – CCGCACGGTGGGTCGCAAT
GlmU 901 reverse - AATCACGGTGTCACGGCCGA
GlmU 1320 reverse – CGTGACCGGCGCAACAAA

The above primers were diluted into 10 µM in RNAse free H_2O and 1 µl added to 250 ng WT DNA and made up to 20 µl with RNase free H_2O, in a 96-well microplate prior to submission to the GSK sequencing group.
The *E. coli* codon optimised, GlmU DNA constructs with either N- or C-terminal purification sequences are listed below:

GlmU N-terminal hexahistidine and TEV cleavage site, DNA sequence:

AGGAGGTAAAACATATGGCACATCATCACCACCACCACGCGGAGAACTTGTAT
TTTCAGGGCATGACATTCCCAGGTGACACAGCAGTTTTAGTATTAGCAGCAGG
TCCAGGTACACGCATGCGCAGCGATACCCCGAAAGTGCTGCATACGTTGGCG
GGTCGCAGCATGTTGAGCCACGTTCTGCACGCGATTGCCAAACTGGCTCCGC
AGCGTTTGATTGTTGTGCTGGGTCACGACCACCAACGTATTGCTCCGCTGGTT

GGCGAGCTGGCGGATACCCTGGGCCGTACGATTGATGTTGCGCTGCAGGAT
CGTCCGCTGGGCACCGGCCATGCGGTCCTGTGCGGCCTGAGCGCGCTGCCG
GATGACTACGCCGGTAACGTCGTGGTGACGTCCGGTGATACCCCGTTGTTGG
ACGCGGACACCCTGGCCGACCTGATTGCGACCCACCGTGCGGTGAGCGCAG
CCGTGACCGTGCTGACTACCACCTTGGACGATCCGTTCGGCTACGGTCGTAT
TCTGCGCACGCAAGACCACGAAGTGATGGCAATCGTGGAGCAGACCGATGC
GACCCCGTCCCAGCGCGAAATCCGTGAGGTCAATGCCGGTGTCTATGCGTTT
GATATCGCAGCTCTGCGTTCTGCTCTGAGCCGTCTGTCCAGCAACAACGCGC
AGCAGGAGCTGTACCTGACCGATGTCATTGCGATCCTGCGTAGCGACGGTCA
AACCGTGCACGCTAGCCATGTCGACGACTCTGCACTGGTCGCGGGTGTGAAC
AATCGTGTTCAGTTGGCGGAGCTGGCATCCGAACTGAATCGCCGCGTTGTGG
CAGCGCACCAACTGGCGGGTGTCACCGTCGTGGACCCGGCGACTACGTGGA
TCGACGTTGACGTGACGATCGGCCGTGACACCGTGATTCACCCGGGTACCCA
GCTGCTGGGTCGTACGCAAATCGGCGGTCGTTGTGTGGTTGGTCCTGATACC
ACTCTGACCGACGTTGCAGTGGGTGACGGTGCGAGCGTTGTACGCACCCAC
GGCTCTAGCAGCAGCATTGGTGATGGTGCAGCCGTTGGCCCGTTTACGTATC
TGCGTCCGGGCACCGCGCTGGGTGCCGATGGTAAGCTGGGTGCGTTCGTCG
AAGTTAAGAATAGCACGATTGGCACCGGCACGAAAGTCCCGCATCTGACCTAT
GTTGGCGATGCAGATATCGGTGAGTACTCGAATATCGGCGCATCCAGCGTTTT
CGTCAACTACGACGGTACGAGCAAGCGTCGCACGACCGTTGGCAGCCATGTG
CGTACTGGCTCTGATACCATGTTTGTTGCGCCGGTCACGATCGGTGATGGCG
CTTATACGGGTGCCGGTACGGTGGTCCGTGAAGATGTCCCGCCAGGTGCGCT
GGCGGTTAGCGCTGGTCCGCAACGCAACATCGAGAACTGGGTTCAGCGTAAA
CGCCCTGGCAGCCCGGCAGCACAAGCCAGCAAGCGTGCGAGCGAAATGGCG
TGCCAGCAGCCGACCCAACCGCCGGACGCCGACCAAACCCCGTAACTCGAG

GlmU C-terminal hexahistidine and TEV cleavage site, DNA sequence:

AGGAGGTAAAACATATGACATTCCCAGGTGACACAGCAGTTTTAGTATTAGCA
GCAGGTCCAGGTACACGCATGCGCAGCGATACCCCGAAAGTGCTGCATACGT
TGGCGGGTCGCAGCATGTTGAGCCACGTTCTGCACGCGATTGCCAAACTGGC
TCCGCAGCGTTTGATTGTTGTGCTGGGTCACGACCACCAACGTATTGCTCCGC
TGGTTGGCGAGCTGGCGGATACCCTGGGCCGTACGATTGATGTTGCGCTGCA
GGATCGTCCGCTGGGCACCGGCCATGCGGTCCTGTGCGGCCTGAGCGCGCT

GCCGGATGACTACGCCGGTAACGTCGTGGTGACGTCCGGTGATACCCCGTTG
TTGGACGCGGACACCCTGGCCGACCTGATTGCGACCCACCGTGCGGTGAGC
GCAGCCGTGACCGTGCTGACTACCACCTTGGACGATCCGTTCGGCTACGGTC
GTATTCTGCGCACGCAAGACCACGAAGTGATGGCAATCGTGGAGCAGACCGA
TGCGACCCCGTCCCAGCGCGAAATCCGTGAGGTCAATGCCGGTGTCTATGCG
TTTGATATCGCAGCTCTGCGTTCTGCTCTGAGCCGTCTGTCCAGCAACAACGC
GCAGCAGGAGCTGTACCTGACCGATGTCATTGCGATCCTGCGTAGCGACGGT
CAAACCGTGCACGCTAGCCATGTCGACGACTCTGCACTGGTCGCGGGTGTGA
ACAATCGTGTTCAGTTGGCGGAGCTGGCATCCGAACTGAATCGCCGCGTTGT
GGCAGCGCACCAACTGGCGGGTGTCACCGTCGTGGACCCGGCGACTACGTG
GATCGACGTTGACGTGACGATCGGCCGTGACACCGTGATTCACCCGGGTACC
CAGCTGCTGGGTCGTACGCAAATCGGCGGTCGTTGTGTGGTTGGTCCTGATA
CCACTCTGACCGACGTTGCAGTGGGTGACGGTGCGAGCGTTGTACGCACCCA
CGGCTCTAGCAGCAGCATTGGTGATGGTGCAGCCGTTGGCCCGTTTACGTAT
CTGCGTCCGGGCACCGCGCTGGGTGCCGATGGTAAGCTGGGTGCGTTCGTC
GAAGTTAAGAATAGCACGATTGGCACCGGCACGAAAGTCCCGCATCTGACCT
ATGTTGGCGATGCAGATATCGGTGAGTACTCGAATATCGGCGCATCCAGCGT
TTTCGTCAACTACGACGGTACGAGCAAGCGTCGCACGACCGTTGGCAGCCAT
GTGCGTACTGGCTCTGATACCATGTTTGTTGCGCCGGTCACGATCGGTGATG
GCGCTTATACGGGTGCCGGTACGGTGGTCCGTGAAGATGTCCCGCCAGGTG
CGCTGGCGGTTAGCGCTGGTCCGCAACGCAACATCGAGAACTGGGTTCAGC
GTAAACGCCCTGGCAGCCCGGCAGCACAAGCCAGCAAGCGTGCGAGCGAAA
TGGCGTGCCAGCAGCCGACCCAACCGCCGGACGCCGACCAAACCCCGGAGA
ATCTGTACTTTCAGGGTGCCCACCATCATCATCACCACGCGTAACTCGAG

The corresponding, predicted protein sequences for GlmU with either N- or C-terminal additions are listed below:

GlmU N-terminal hexahistidine and TEV cleavage site predicted protein sequence:

MAHHHHHHAE NLYFQGMTFP GDTAVLVLAA GPGTRMRSDT PKVLHTLAGR SMLSHVLHAI AKLAPQRLIV
VLGHDHQRIA PLVGELADTL GRTIDVALQD RPLGTGHAVL CGLSALPDDY AGNVVVTSGD TPLLDADTLA
DLIATHRAVS AAVTVLTTTL DDPFGYGRIL RTQDHEVMAI VEQTDATPSQ REIREVNAGV YAFDIAALRS
ALSRLSSNNA QQELYLTDVI AILRSDGQTV HASHVDDSAL VAGVNNRVQL AELASELNRR VVAAHQLAGV
TVVDPATTWI DVDVTIGRDT VIHPGTQLLG RTQIGGRCVV GPDTTLTDVA VGDGASVVRT HGSSSSIGDG
AAVGPFTYLR PGTALGADGK LGAFVEVKNS TIGTGTKVPH LTYVGDADIG EYSNIGASSV FVNYDGTSKR
RTTVGSHVRT GSDTMFVAPV TIGDGAYTGA GTVVREDVPP GALAVSAGPQ RNIENWVQRK RPGSPAAQAS
KRASEMACQQ PTQPPDADQT P

GlmU C-terminal hexahistidine and TEV cleavage site predicted protein sequence:

```
MTFPGDTAVL VLAAGPGTRM RSDTPKVLHT LAGRSMLSHV LHAIAKLAPQ RLIVVLGHDH QRIAPLVGEL
ADTLGRTIDV ALQDRPLGTG HAVLCGLSAL PDDYAGNVVV TSGDTPLLDA DTLADLIATH RAVSAAVTVL
TTTLDDPFGY GRILRTQDHE VMAIVEQTDA TPSQREIREV NAGVYAFDIA ALRSALSRLS SNNAQQELYL
TDVIAILRSD GQTVHASHVD DSALVAGVNN RVQLAELASE LNRRVVAAHQ LAGVTVVDPA TTWIDVDVTI
GRDTVIHPGT QLLGRTQIGG RCVVGPDTTL TDVAVGDGAS VVRTHGSSSS IGDGAAVGPF TYLRPGTALG
ADGKLGAFVE VKNSTIGTGT KVPHLTYVGD ADIGEYSNIG ASSVFVNYDG TSKRRTTVGS HVRTGSDTMF
VAPVTIGDGA YTGAGTVVRE DVPPGALAVS AGPQRNIENW VQRKRPGSPA AQASKRASEM ACQQPTQPPD
ADQTPENLYF QGAHHHHHHA
```

2.2.3 GlmU wild type and H374A expression

GlmU wild type and H374A expression was carried out by using 1 µl of the desired plasmid at 100 ng/µl to transform BL21(DE3)pLysS *E. coli* competent cells (Thermo Fisher Scientific cat. no. C606003) at 25 °C for 10 mins, followed by a heat shock step at 42 °C for 45 secs and then a further incubation step at 4 °C for 2 mins. The transformants were then plated on kanamycin and 1% glucose containing LB agar plates and incubated for 12 hours at 37 °C. Following the overnight incubation, two colonies per plate were picked and used to inoculate two 200 ml flasks each filled with 100 ml LB broth containing kanamycin and 1% glycerol.

GlmU wild type protein expression was carried out in BL21(DE3)pLysS cells, grown in LB broth, supplemented with 1% glucose and 0.05 mg/ml kanamycin, at 37 °C until an optical density at 600 nm of 0.5 – 0.8 was achieved. The temperature was subsequently shifted to 25 ºC and the culture induced with 1 mM IPTG and then incubated for a further 20 hrs at the National Institute for Medical Research (NIMR) Large Scale Laboratory, using a 20-litre fermenter. A total yield of 120 g of cell pellet were obtained in an overnight fermentation.

GlmU H374A cultures were incubated at 30 °C, while shaking at 200 rpm for 16 hours before adding 20 mls to 2.5 L ultra-flasks, which contained 1 L of Turbo medium, plus 1 ml of 50 mg/ml kanamycin, 1% glucose and 0.5 ml Antifoam. The ultra-flasks were then incubated at 37 °C, while shaking at 200 rpm until an optical density at 600 nm of 0.5 – 0.8 was achieved and then inducing with the addition of 1 mM IPTG. The absorbance of the induced cells was monitored until an optical density at 600 nm of 5.5 was achieved. The cultures were then harvested by centrifugation at 4000 x g for 20 mins while maintaining a temperature of 4 °C.

2.2.4 GlmU WT protein purification

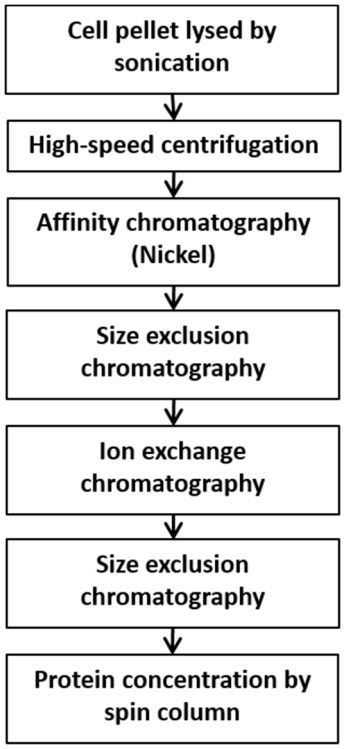


Figure 17. GlmU WT protein purification workflow.

This protein purification workflow was applied to four separate protein preparations; one *C*-terminal hexahistidine GlmU, one *N*-terminal hexahistidine GlmU in the absence of 10 mM $MgCl_2$ and two *N*-terminal hexahistidine GlmU in the presence of 10 mM $MgCl_2$. Details of the methods at each stage can be found in the method section below.

Throughout the GlmU protein purification process, all steps were performed at 4 °C. Frozen BL21(DE3)pLysS (*pJ411*::GlmU) cells were thawed on ice, and lysed by sonication, in the presence of lysis buffer, either [25 mM N-[2-hydroxyethyl]piperazine-N′-[2-ethanesulfonic acid] (HEPES) at pH 7.5, 100 mM NaCl and 1 mM EDTA], or [25 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$ and 1 mM EDTA], with both buffers supplemented with 10 mg/ml lysozyme, 1 µl/ml benzonase and 1 ml/L protease inhibitor cocktail (Sigma Aldrich cat. no. P8340). Soluble protein was separated from cell debris by ultracentrifugation at 100000 x *g* for 60 min. The soluble fraction was loaded on to two in-line, 5 ml Ni-NTA columns (GE Healthcare HisTRAP HP, cat. no. 17-5248-02) pre-equilibrated with 10 column volumes at 5 ml min^{-1} of buffer A, either [25 mM HEPES (pH 7.5), 100 mM NaCl and 10 mM Imidazole], or [25 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$ and 10

mM Imidazole]. The hexahistidine tagged protein was loaded onto the Ni-NTA column at 1 ml min^{-1}, followed by a column wash with 5 column volumes of buffer A. The sample is then eluted at 1 ml min^{-1} using a two-phase gradient of buffer B, either [25 mM HEPES (pH 7.5), 100 mM NaCl and 1 M Imidazole] or [25 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$ and 1 M Imidazole], with the first phase from 0 % to 50% buffer B in 20 column volumes and the second phase, from 51 % to 100 % buffer B covered in 10 column volumes. Fractions of 2 ml volumes were collected in high volume, polypropylene, 96-well microplates (Nunc cat. no. 278743) and the samples spanning the ÄKTA chromatograph $A_{280\ nm}$ peak, were then analysed by SDS−PAGE gel electrophoresis. The fractions containing only a single GlmU-sized band were pooled together and loaded onto a HiTrap SP HP 5 ml ion exchange column (GE Healthcare, cat. no. 17115201), pre-equilibrated with 30 column volumes of buffer C, either [25 mM HEPES (pH 7.5)], or [25 mM HEPES (pH 7.5), 10 mM $MgCl_2$] at 1 ml min^{-1}. The protein was loaded onto the column at 0.5 ml min^{-1} and subsequently eluted at 0.5 ml min^{-1} using a gradient from 0 % to 100 %, over 20 column volumes of buffer D, either [25 mM HEPES (pH 7.5), 1 M NaCl] or [25 mM HEPES (pH 7.5), 10 mM $MgCl_2$ and 1 M NaCl]. The fractions spanning the chromatograph $A_{280\ nm}$ peak, were then analysed by SDS−PAGE gel electrophoresis and only those containing a single GlmU-sized band were selected and combined for further purification. The pooled protein was loaded onto a 320 ml HiLoad 26/60 Superdex 200 prep grade column (GE Healthcare, cat. no. 17-1071-01) at 1 ml min^{-1} using a 10 ml Superloop, which had previously been pre-equilibrated with 1.2 column volumes of buffer E, either [25 mM HEPES (pH 7.5), 100 mM NaCl] or [25 mM HEPES (pH 7.5), 10 mM $MgCl_2$ and 100 mM NaCl]. The protein sample was then eluted at 1 ml min^{-1} over 1.2 column volumes using buffer E and 2 ml fractions were collected in high volume, polypropylene, 96-well microplates (Nunc cat. no. 278743). Fractions were selected that spanned the ÄKTA chromatograph $A_{280\ nm}$ peak and analysed by SDS-PAGE gel electrophoresis, with only the samples that contained GlmU-sized, single bands combined. In some cases, particularly for structural biology studies, it was necessary to remove the hexahistidine tag from a proportion of the semi-purified material, which was separated after the ion exchange step. An aliquot of 10 mls of protein solution was combined with 100 µl of 10 mg/ml TEV protease and incubated at 4 °C for 12 hours, while the solution was continuously

mixed, before the aforementioned size exclusion chromatography process was followed.

2.2.5 Site directed mutagenesis

The following mutations, GlmU R344A, GlmU K362A, GlmU H374A, GlmU Y377A and GlmU N388A, were attempted using site-directed mutagenesis (SDM) of the WT expression construct. The following mutagenesis primers were designed following the instructions detailed in the Phusion (Thermo Scientific, cat. no. F-541) SDM method, synthesised by Integrated DNA Technologies (IDT) and subsequently prepared as 25 mM stocks in RNAse free H_2O:

Primer sequences
R344A forward - CCGTTTACGTATCTGgccCCGGGCACCGCGCTG
R344A reverse - GCCAACGGCTGCACCATCACCAATGCTGCTGCT
K362A forward - GCGTTCGTCGAAGTTgccAATAGCACGATTGGC
K362A reverse - ACCCAGCTTACCATCGGCACCCAGCGCGGTGCC
H374A forward – GGCACGAAAGTCCCGgccCTGACCTATGTTGGC
H374A reverse – TTCGTCGAAGTTAAGAATAGCACGATTGGCACC
Y377A forward - AGTCCCGCATCTGACCgccGTTGGCGATGCAGA
Y377A reverse – TTCGTGCCGGTGCCAATCGTGCTATTCTTAACT
N388A forward - ATCGGTGAGTACTCGgccATCGGCGCATCCAGC
N388A reverse - ATCTGCATCGCCAACATAGGTCAGATGCGGGAC

Linear amplification for each mutant was carried out by adding 32 ng of WT plasmid (stock concentration of 3.2 ng/µl) using 0.5 µM of phosphorylated SDM primer pairs and 0.02 U/µl Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Scientific) into a total volume of 50 µl. The PCR cycling protocol, carried out on a BioRad C1000 Touch Thermal Cycler (BioRad), was one cycle of an initial denaturation at 98 °C for 30 s, followed by 25 cycles consisting of denaturation at 98 °C for 5 s, annealing at 68 °C for 30 s and extension for 72 °C 30 s/kb. The final extension step is carried out at 72 °C for 10 mins followed by an indefinite hold at 4 °C.
PCR success was assessed by agarose gel electrophoresis, which was carried out by adding 5 µl of each of the PCR mixes to a E-Gel with SYBR Safe (Thermo

Scientific) and the samples separated using the E-Gel iBase Power system (Thermo Scientific) followed by image capture on a G-Box imager (Syngene).

The methylated WT template DNA was then degraded using an addition of 1 µl DpnI (New England Biolabs, cat. no. R0176S) to the PCR mixes, followed by an incubation at 37 °C for 60 mins.

Ligation reactions were carried out using 20 ng PCR product that was circularised using 0.5 µl T4 DNA Ligase in a 10 µl reaction volume for 5 mins at 25 °C, followed by storage at 4 °C.

Ligated SDM DNA was then used for transformation of One Shot chemically competent TOP10 *E. coli* cells (Thermo Scientific cat. no. C404006). After addition of 1 µl each of SDM DNA to 50 µl TOP10 *E. coli* cells, there was an incubation of 10 mins followed by a heat shock step at 42 °C for 45 s and then a further incubation step at 4 °C for 2 mins. The cells were then added to 250 µl SOC medium and incubated at 37 °C, while shaking, for one hour before adding 100 µl of competent cells to Kanamycin containing LB agar plates. The agar plates were then incubated for 12 hours at 37 °C before picking two colonies per SDM mutant and adding to two separate 200 ml LB broths containing Kanamycin. The cells were incubated, with vigorous shaking, for 12 hours at 37 °C and then harvested by centrifugation at 8000 rpm for 3 mins at 25 °C.

SDM DNA was then amplified using the Qaigen Miniprep kit following the QIAprep Spin Miniprep and microcentrifuge protocol (Qaigen cat. no. 27106).

The DNA concentration of the SDM preparations was assessed using the NanoDrop spectrophotomer (Thermo Scientific) by adding 1 µl and measuring the absorbance of nucleic acids at 260 nm.

To confirm SDM DNA sequences, Sanger sequencing was used, utilising the following primers which were previously used for WT DNA sequence confirmation:

Sequencing primer sequences
GlmU 901 forward - TCGGCCGTGACACCGTGATT
GlmU 1320 reverse - CGTGACCGGCGCAACAAA

The primers were then diluted into 10 µM in RNAse free H_2O and 1 µl added to 250 ng SDM DNA and made up to 20 µl with RNase free H_2O, in a 96-well microplate prior to submission to the GSK sequencing group.

2.2.6 GlmU H374A purification

The protein purification of GlmU H374A was carried out following a modified version of the protocol used for the WT enzyme.

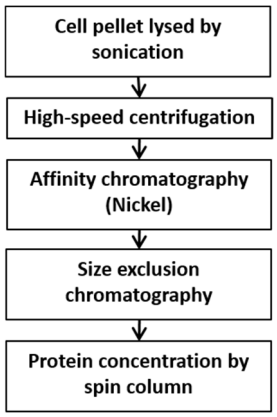


Figure 18. GlmU H374A protein purification workflow.

This protein purification workflow was applied to two separate protein preparations of *N*-terminal hexahistidine GlmU H374A in the presence of 10 mM MgCl2. Details of the methods at each stage can be found in the method section below. The ion exchange and additional SEC steps (Figure 17) were omitted as the GlmU H374A protein, post-SEC (1), had no contaminants.

Frozen BL21(DE3)pLysS (*pJ411*::GlmU H374A) cells were thawed on ice, and lysed by sonication, in the presence of lysis buffer consisting of 25 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM MgCl2 and 1 mM EDTA, supplemented with 10 mg/ml lysozyme, 1 µl/ml benzonase and 1 ml/L protease inhibitor cocktail (Sigma Aldrich cat. no. P8340). Soluble protein was obtained by centrifugation at 100000 x *g* at 4 °C for 60 min. The soluble fraction was loaded on to two in-line, 5 mL Ni-NTA columns, pre-equilibrated with 10 column volumes of buffer A comprising of 25 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM MgCl2 and 10 mM Imidazole, at 3.5 ml min^{-1}. The hexahistidine tagged protein was loaded onto the column at 1 ml min^{-1} and then eluted using the same gradient as previously described for the WT protein of buffer B comprising of 25 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM MgCl2 and 1 M Imidazole. Fractions spanning the ÄKTA chromatograph $A_{280\ nm}$ peak were analysed by SDS−PAGE. Fractions containing only a single GlmU-sized band were pooled together and loaded onto a 320 ml HiLoad 26/60 Superdex 200 prep grade column, pre-equilibrated with size exclusion chromatography (SEC) buffer, comprising of 25 mM HEPES (pH 7.5), 10 mM MgCl2 and 100 mM NaCl. The protein was eluted at 1 ml min^{-1} over 1.2 column volumes using SEC buffer. Peak fractions

were analysed by SDS-PAGE and only those containing GlmU were pooled together. Hexahistidine tag removal was carried out on a proportion of the purified material, which was separated after the affinity chromatography step. GlmU was incubated with TEV protease at 4 °C overnight and then was loaded onto a 320 ml HiLoad 26/60 Superdex 200 prep grade column, prior to following the previously described SEC elution protocol. Purified GlmU H374A was analysed by SDS-PAGE gel electrophoresis and electrospray ionization mass spectrometry (ESI-MS) for accurate determination of the monomer mass.

2.2.7 Further GlmU SDM expression and purification

Expression and purification of the SDM mutants GlmU R344A, GlmU K362A and GlmU Y377A was carried out using a high throughput, magnetic bead-based methodology developed at GSK. To control for use of a novel protein purification process, GlmU WT and GlmU H374A were prepared to allow comparison with material prepared by the previously described column chromatography-based methods. Frozen BL21(DE3) pLysS cells were thawed on ice and transferred, in 12 µl volume, to five wells of a 96-well PCR microplate, followed by addition of 50-100 ng of each of the SDM plasmids before leaving at 4 °C for 30 mins. The 96-wel PCR microplate was then transferred to a PCR block and subjected to 42 °C, to heat shock the cells for 1 min before putting the microplate back on ice for 2 mins. The cells were then transferred to wells containing 500 µl LB media in a high volume, polypropylene, 96-well microplate block (Nunc cat. no. 278743) and incubated for 37 °C, while shaking at 200 rpm for 2 hrs. Following the incubation, 500 µl LB media, containing 50 µg/ml Kanamycin, was added to each cell suspension, before incubation at 37 °C, while shaking at 200 rpm, for 12 hrs.

A high volume, polypropylene, 96-well microplate block (Nunc cat. no. 278743), was filled with 1 ml of Overnight Express media per active well, supplemented with 50 µg/ml Kanamycin, before inoculating 8 wells per plasmid with 30 µl of the overnight seed cultures (total of 40 wells inoculated). The 96-well block was then incubated for 37 °C, while shaking at 200 rpm, for 2 hrs, before reducing the incubation temperature to 17 °C and increasing the shaking speed to 300 rpm, for 12 hrs.

After overnight incubation, the 8 x 1 ml cultures for each construct were collated into one well of a 24-well microplate, before centrifugation at 4000 rpm for 30 mins followed by removal of the supernatant and freezing of the pellets for 12 hrs at -20 °C. After overnight incubation, the 24-well microplate was thawed, and 700 µl of lysis buffer was added to each pellet and pipetted to resuspend before incubation at room temperature, while shaking at 120 rpm for 2 hrs. The 24-well microplate was then centrifuged for 10 mins at 3000 rpm to prepare clarified lysate for purification using the KingFisher™ Flex System (Thermo Fisher Scientific cat. no. 5400610).

Following centrifugation, 150 µl of Pierce™ Ni-NTA Magnetic Agarose Beads (Thermo Fisher Scientific cat. no. 78605) to 5 wells, one per construct, on a KingFisher™ polypropylene, V-bottom, high volume 96-well microplate (Thermo Fisher Scientific cat. no. 95040450). These additions were , followed by a 350 µl addition of equilibration buffer [50 mM sodium phosphate, 0.3 M NaCl, 10 mM imidazole, 0.05% Tween™-20, pH 8.0]. In addition, in a further high volume 96-well microplate, 500 µl of equilibration buffer was added to every well. Furthermore, another 96-well microplate, was prepared by adding 700 µl of each construct to separate wells, followed by additions of 300 µl equilibration buffer to each condition. Two 96-well microplates were prepared to act wash plates, by adding 1 ml of wash buffer [50 mM sodium phosphate, 0.3 M NaCl, 15 mM imidazole, 0.05% Tween™-20, pH 8.0] in every well. Finally, two high volume 96-well microplates were filled with 100 µl elution buffer [50 mM sodium phosphate, 0.3 M NaCl, 0.3 M imidazole pH 8.0]. All the microplates were then added to the stage of the KingFisher Flex System and purification programme initiated using the following basic protocol; cells were lysed, followed by collection of magnetic beads, which were then added to the aforementioned sample. The beads were agitated and then collected, before washing of the samples, prior to elution of the protein and removal of the beads.

The two elution plates were removed and 100 µl per sample collated into the same well on one 96-well microplate, which are then further collated to combine all the individual construct wells into one well per SDM or WT enzyme.

To minimise the amount of imidazole present in the buffer of the pooled proteins, the sample were concentrated by centrifugation, using Amicon Ultra centrifuge tubes (see Section 2.2.1) and diluted in storage buffer [25 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$]. This process of concentration and subsequent dilution in storage buffer was repeated 5 times, prior to snap freezing of the samples on dry ice.

The five GlmU protein preparations were then assessed using SDS-PAGE gel electrophoresis (see Section 2.2.1) and Caliper LabChip GXII Touch High Throughput (HT) Protein Characterisation System (Perkin Elmer cat. no. CLS138161). The Caliper LabChip instrument utilises microfluidics to separate the denatured protein or molecular weight samples, utilising the following basic methodology; samples were denatured by the addition of 1 M DTT before being added to a 96-well microplate. The samples and ladder were further denatured by application of 100 °C for 5 mins, before adding 32 µl to the test samples and 120 µl to the molecular weight ladder marker. The test microplate was taken into the instrument and the samples and marker were then aspirated, separated and detected by the LabChip GXII.

2.2.8 Measurement of acetyltransferase activity

GlmU acetyltransferase forward reaction was monitored by reacting 4,4′ dithiodipyridine (DTP) with the free thiol of the product coenzyme A (CoA-SH), which leads to the formation of 4-thiopyridone (4-TP) (Scheme 7). The formation of 4-TP causes a change in absorbance at 324 nm (ε = 19800 M^{-1} cm^{-1})[392,393], which leads to an increase in optical density (OD). A typical reaction mix contained 50 mM HEPES at pH 7.5, 50 mM NaCl, 10 mM MgCl2, 1 mM 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (CHAPS), 400 µM DTP, 100 µM Ac-CoA and 200 µM GlcN-1P. All reagents, when prepared as intermediate stocks in the aforementioned assay buffer, were stored at 4 °C prior to use. Reactions were carried out in 100 µl, in a black, polystyrene 384-well, clear bottomed microplate and were generally initiated by an addition of GlmU, typically 5 nM final concentration, unless otherwise stated. The microplate containing the reaction mixture was mixed by centrifugation at 1000 rpm for 1 minute, prior to the final, initiating reagent addition. Following the reaction initiating addition, the microplate was immediately transferred to a Tecan M1000 PRO or Safire2 monochromator multimode readers and the change in absorbance monitored at 324 nm ± 5 nm, using 5 flashes per well, at 30 °C. The microplate was monitored for a duration of 30 to 45 minutes, using the minimum time interval, determined by the number of wells being read. Kinetic data analysis was carried out as detailed in Figure 19.

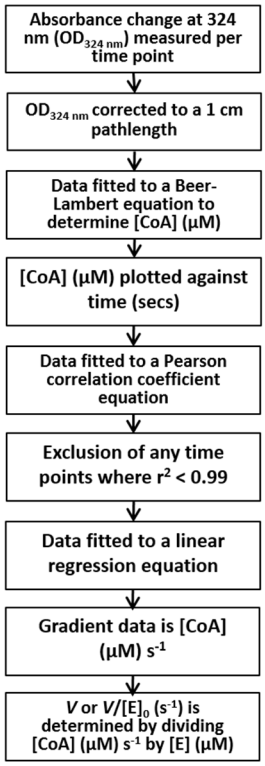
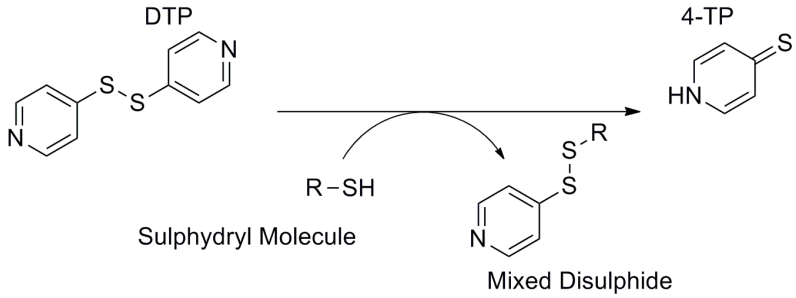


Figure 19. Acetyltransferase absorbance assay data analysis workflow.

The workflow presented above is specific for the use of DTP. When using DTNB absorbance at OD_{412nm} is measured but all other data analysis steps remain the same.

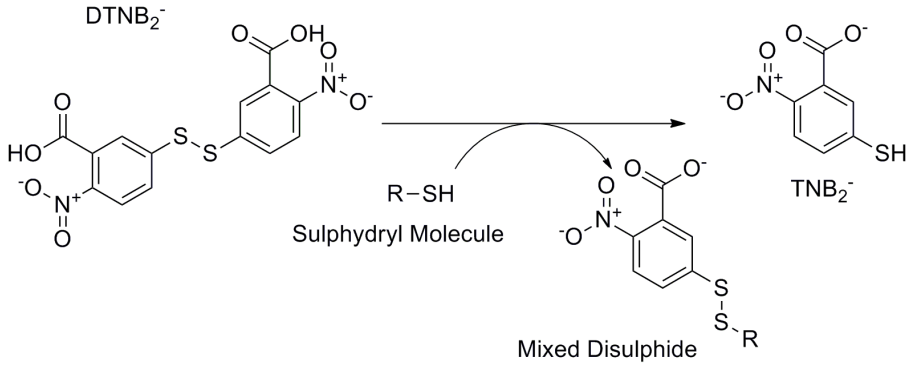
Absorbance change at 324 nm, measured as optical density, was pathlength corrected to 1 cm and then converted to concentration of CoA-SH formed by utilising the absorbance extinction coefficient of 4-TP and Beer-Lambert law (see Section 2.2.30)[394]. Subsequently, the concentration of CoA-SH per time point were plotted against time (in secs) and fitted to a Pearson correlation coefficient equation (see Section 2.2.30). All data with a r^2 value greater than 0.99 were then fitted to a linear regression equation to determine both the gradient and intercept (see Section 2.2.30).



Scheme 7 DTP detection of sulphydryl molecules.

Reaction scheme for DTP detection of the sulphydryl of CoA-SH.

An alternative thiol detection approach, utilising 5,5-Dithiobis(2-nitrobenzoic acid) ($DTNB^{2-}$, or DTNB)[395] were investigated to monitor the GlmU acetyltransferase forward reaction. DTNB reacts with the free thiol of CoA-SH that leads to formation of 2-nitro-5-thiobenzoic acid (NTB^{2-}), which is photometrically detectable, with a maximum absorption at 412 nm (ε = 14140 M^{-1} cm^{-1}). The assay protocol for using the alternative thiol, absorbance detection is unchanged, other than replacing 400 µM DTP, with equimolar amount of DTNB (Scheme 8).



Scheme 8 DTNB detection of sulphydryl molecules.
Reaction scheme for DTNB detection of the sulphydryl of CoA-SH.

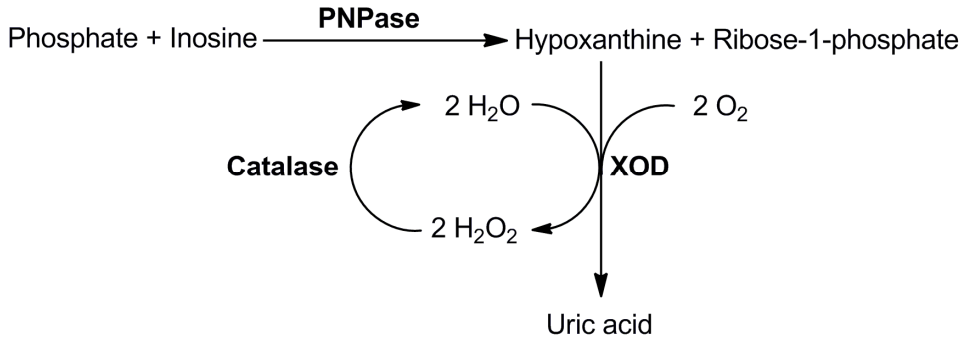
To monitor sensitivity of the thiol detection approaches, titrations of CoA-SH were carried out from 500 µM, in the standard assay buffer, using fixed, previously stated concentrations of either DTP or DTNB and the corresponding microplate. The CoA-SH titration was mixed with the desired detection reagent by centrifugation at 1000 rpm for 1 minute, prior to measurement of the assay signal using either the Tecan M1000 PRO or Safire2 monochromator multimode microplate readers.

2.2.9 Preparation of uridylyltransferase reagents

Reagents used for GlmU uridylyltransferase studies had to have low free inorganic phosphate (Pi) contamination, as both assay methodologies utilised detect an increase in Pi, which are coupled through pyrophosphatase (PPase) to convert the product, pyrophosphate (PPi), to Pi. Any contaminating Pi in the assay reagents will increase the background of the assay and subsequently will minimise the maximum attainable signal window.

To minimise the concentration of contaminating Pi, reagents were treated with a coupled enzyme system to utilise free Pi and convert it to either ribose-1-phosphate,

uric acid or H_2O. Scheme 9 shows the coupled system utilised to prepare reagents for uridylyltransferase studies.



Scheme 9. Reaction scheme of enzyme coupled Pi scrubbing.
Reaction scheme for Pi removal enzyme coupled system.

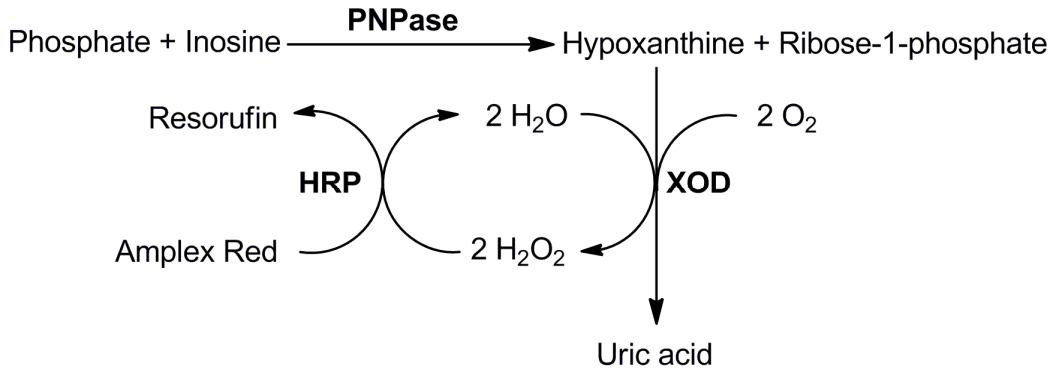
Initially, to remove potential contaminants, 200 µl catalase immobilised on agarose beads was added to 1 ml deionised (DI) H_2O in 1.5 ml microcentrifuge tubes and mixed thoroughly by inversion before centrifugation at 2500 rpm for 2 minutes using a 5418R bench top centrifuge (Eppendorf, cat. no. 5418000068) and removal of the supernatant. This process was repeated two times, then after the final supernatant removal, 200 µl immobilised catalase was incubated at 4 °C while the reagent mix was prepared. A typical reaction mix comprised of 1 ml assay buffer [50 mM HEPES (pH 7.4), 50 mM NaCl, 10 mM $MgCl_2$] was prepared, which contained the following reagents; 10 mM reagent to be scrubbed (for example, UTP or GlcNAc-1P), 22.5 mM Inosine, 4 U/ml Xanthine oxidase (XOD) and 0.3 U/ml Purine nucleoside phosphorylase (PNPase). The reaction mix was then added in 1 ml volume to 200 µl immobilised catalase and subsequently mixed at 4 °C on a tube roller for 30 minutes. After the incubation, the mixture was then centrifuged for 1 min at 13000 rpm to pellet the immobilised catalase and the supernatant reagent was removed, sub-aliquoted and stored at -20 °C.

2.2.10 Measurement of uridylyltransferase activity

Initial velocities for the forward uridylyltransferase reaction of GlmU were performed at 30 °C using inorganic pyrophosphatase (PPase) and two different assay methodologies to detect the formation of pyrophosphate (PPi) and subsequently inorganic phosphate (Pi). The two assay methodologies have both been previously described and are; an enzyme coupled system to detect the formation of free

phosphate in solution through the formation of the fluorescent product resorufin[396], and a 7-Diethylamino-3-((((2-Maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC)-labelled phosphate binding protein (MDCC-PBP)[397].

The enzyme coupled phosphate detection assay (referred to hereon in as PiXan) relies on coupling phosphate generation to purine nucleoside phosphorylase, xanthine oxidase, and horseradish peroxidase, which is detailed in Scheme 10. The assay utilised the following buffer; [50 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM CHAPS], with care taken to minimise buffer phosphate contamination by using MQ DI H_2O and avoiding re-used glassware. A typical reaction mix was prepared in the assay buffer and contained the following; 1.5 mM inosine, 50 µM Amplex Red, 0.02 IU/ml PNPase, 0.4 IU/ml XOD, 1 IU/ml horseradish peroxidase (type XII) (HRP), 20 µM UTP and 50 µM GlcNAc-1P. The fluorescence signal obtained was measured using a Tecan Infinite M1000 Pro microplate reader using a wavelength of 530 nm for the excitation and 590 nm for the emission. Reactions were initiated by the addition of enzyme, typically at final concentrations of 5 nM GlmU and 1 IU/ml PPase.



Scheme 10. Reaction scheme of enzyme coupled Pi detection assay (PiXan). Reaction scheme for the enzyme coupled assay utilised to detect Pi.

The MDCC-PBP biosensor assay is based on binding of Pi to the A197C mutant of *E. coli* phosphate binding protein (PBP), which causes a conformational change in the biosensor, alleviating fluorophore quenching, leading to an increase in fluorescence intensity. A typical reaction mix was prepared in assay buffer [50 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM MgCl2, 1 mM CHAPS] and contained 15 µM PBP, 20 µM UTP and 50 µM GlcNAc-1P. The fluorescence signal obtained was measured using a Tecan Infinite M1000 Pro microplate reader exciting the MDCC fluorophore of PBP at a wavelength of 425 nm and measuring fluorescence emission

at 474 nm. Reactions were initiated by the addition of enzyme, typically at final concentrations of 2.5 nM GlmU and 1 IU/ml PPase.

To monitor sensitivity of the Pi detection assays, titrations of a Pi standard were carried out from 500 µM, in the standard assay buffer, using fixed, previously stated conditions for either the enzyme-coupled or the MDCC-PBP approaches and the corresponding microplate. The Pi standard titration was then mixed with the detection assay by centrifugation at 1000 rpm for 1 minute, prior to measurement of the assay signal using either the Tecan M1000 PRO or Safire2 monochromator multimode microplate readers.

2.2.11 Rapidfire Mass Spectroscopy

An alternative GlmU acetyltransferase assay was initially investigated using the RapidFire high-throughput (HT) mass spectrometry system to simultaneously detect decrease in concentrations of GlcN-1P and increase in concentrations of GlcNAc-1P. Initially, GlmU was titrated from 250 nM in an assay buffer [50 mM 2-(N-morpholino)ethanesulfonic acid (MES), 50 mM HEPES, 100 mM ethanolamine, 10 mM $MgCl_2$, and 1 mM CHAPS] in clear, 384-well polypropylene microplates (Greiner, cat. no. 781281) before initiating addition of fixed concentrations of 200 µM GlcN-1P and 100 µM Ac-CoA and then additions of either 0.5 % formic acid to quench the reaction every 1 min. Further small molecule inhibitor screening was carried out using a modification of the aforementioned method maintaining GlmU at a fixed concentration of 4 nM, which was pre-incubated with 100 µM of test compound, prior to reaction initiation with the acetyltransferase substrates.

To help monitor the performance of the Rapidfire MS detection technology, GlcN-1P and GlcNAc-1P were titrated the same assay buffer in the same microplates before addition of 0.5 % formic acid, prior to quantification on the Rapidfire system.

After addition of the acidic quench, microplates were stable and could either be stored at −20 °C or analysed immediately. All reagent additions to the assay microplates were performed using either a Multidrop Combi (ThermoFisher) or dragonfly discovery (TTP Labtech).

Test microplates were transferred onto a RapidFire200 high-throughput autosampler solid-phase extraction (SPE) system (Agilent Technologies) and 10 µl of sample was aspirated directly and loaded onto the RapidFire micro-scale solid-phase C4

extraction cartridge to remove buffer salts in combination with a 2 sec wash cycle, comprised of a HPLC-grade H_2O wash containing 0.1 % formic acid. Analytes were then co-eluted into the mass spectrometer, using a 2.5 sec elution cycle using a solution comprised of 80 % acetonitrile and 0.1 % formic acid. GlcN-1P and GlcNAc-1P were monitored on a Sciex API QqQ mass spectrometer (Applied Biosystems) in negative electrospray injection (ESI) mode following multiple-reaction monitoring (MRM) transitions at 279.3/179.0 and 351.4/271.1, respectively. The mass spectrometer used an ESI voltage of 5500 V and a source temperature of 600 °C, with a dwell time of 50 ms for each transition.

Individual MRM transitions were saved as text files, and the extracted ion chromatograms (XIC) were integrated and processed using the RapidFire peak integration software (version 3.6). The integrated peak area data for each condition monitored were used to determine the extent of conversion of GlcN-1P to GlcNAc-1P (see Section 2.2.30).

2.2.12 Electrospray Injection Liquid Chromatography Mass Spectroscopy

The accurate mass of protein samples was determined using ESI LC-MS on either a Waters 2795 LCT Premier or Agilent 6224 Time-of-flight (TOF) LC-MS equipped with an Agilent 1200 series autosampler, which have the detection limits set at either 800 - 2500 or 1700 - 3200 m/z. For the Waters 2795 LCT Premier, samples were loaded at 0.25 ml min^{-1} onto a 2.1 x 150 mm PLRP-S 8 µm 1000A column (Polymer Labs), which was maintained at a temperature of 60 °C for the duration of the experiment. A gradient was then applied to elute the sample, using buffer A [95 % pure water, 5% acetonitrile, 0.05 % TFA] and buffer B [95 % acetonitrile, 5 % pure water, 0.05 % TFA]. The ESI spectrum was obtained in the multichannel acquisition mode, with scanning from 500 to 1800 m/z at a scan time of 7 secs. The mass spectrometer is equipped with MassLynx and Transform software for data acquisition and spectrum handling.

For the Agilent 6224, 10 µl samples were loaded at 0.5 ml min^{-1} onto a 1 x 50 mm PLRP-S 5 µm 1000A column (Polymer Labs), which was maintained at a temperature of 70 °C for the duration of the experiment. A gradient was then applied to elute the sample, using buffer A [99.8 % pure water, 0.2 % (v/v) formic acid] and buffer B [99.8 % acetonitrile, 0.2 % (v/v) formic acid]. Gradient conditions were

initially 10 % B, increasing to 30 % after 0.5 min, and then linearly to 80 % B over 4.5 min, prior to a 1.2 min flush with 100 % B and then a 1.9 min equilibration with 10 % B prior to the next injection. Mass spectra were then deconvoluted using Agilent MassHunter Qualitative Analysis version B.06.00 over the mass range 35 - 170 kDa to obtain the intact protein mass.

2.2.13 Preparation and quantification of GlmU protein for metal ion modulation studies

To investigate the metal ion dependence of GlmU, initially, two different protein preparations, purified in the absence and presence of Mg^{2+}, were incubated with 25 mM EDTA for 18 hours at 4 °C. To remove EDTA, protein mixtures were dialysed on two occasions for 12 hours, while gently stirring on a magnetic stirrer, at 4 °C in 1 litre of buffer composed of 25 mM HEPES (pH 7.5), 100 mM NaCl using a D-Tube Dialyzer Midi, MWCO 3.5 kDa (Merck Millipore, cat. no. 71506).

The concentration of dialysed GlmU protein preparations were then assessed using a Bradford protein assay (see Section 2.2.4).

Acetyltransferase activity of the dialysed GlmU preparations were assessed by titrating the enzymes and monitoring a CoA-SH release using the absorbance assay (see Section 2.2.8) in a buffer composed of 50 mM HEPES (pH 7.5), 50 mM NaCl, 1 mM CHAPS and either 10 mM $MgCl_2$ or $MnCl_2$, in the presence of 200 µM Ac-CoA, GlcN-1P and 400 µM DTP.

2.2.14 Acetyltransferase metal ion modulation

Metal ion dependence of GlmU acetyltransferase was initially investigated using the CoA-SH release absorbance assay (see Section 2.2.8) by titrating either Ac-CoA or GlcN-1P from 2 mM, at several fixed concentrations of either $MgCl_2$, $MnCl_2$ or $CaCl_2$, ranging from 30 mM to 1.5 mM. Acetyltransferase initial velocities were calculated for each condition and the data fitted to the Michaelis-Menten kinetic model (see Section 2.2.30).

Metal ion effects were investigated by titrating $MgCl_2$, $MnCl_2$, $CaCl_2$, $CoCl_2$, $NiCl_2$, $ZnCl_2$, or $Zn(OAc)_2$ from 30 mM by monitoring the acetyltransferase activity of dialysed preparations of GlmU, in an assay buffer composed of 50 mM HEPES (pH 7.5), 50 mM NaCl, 1 mM CHAPS, in the presence of 200 µM Ac-CoA, GlcN-1P and

400 µM DTP. Acetyltransferase initial velocities were calculated for each condition and the data fitted, where appropriate, to either Michaelis-Menten kinetic or four parameter IC_{50} model (see Section 2.2.30).

Further metal ion effects as competitors to Mg^{2+}, either in the assay buffer or bound to the enzyme, were investigated by titrating $MgCl_2$, $MnCl_2$, $CaCl_2$, $CoCl_2$, $NiCl_2$, $ZnCl_2$, or $Zn(OAc)_2$ from 30 mM by monitoring the acetyltransferase activity of dialysed or non-dialysed GlmU, in an assay buffer composed of 50 mM HEPES (pH 7.5), 50 mM NaCl, 1 mM CHAPS with or without 10 mM $MgCl_2$, in the presence of 200 µM Ac-CoA, GlcN-1P and 400 µM DTP. Acetyltransferase initial velocities were calculated for each condition and the data fitted, where appropriate, to either Michaelis-Menten kinetic or four parameter IC_{50} model (see Section 2.2.30).

2.2.15 Acetyltransferase initial velocity patterns

Initial velocity pattern experiments were carried out for GlmU acetyltransferase using the CoA-SH release absorbance assay (see Section 2.2.8) by titrating Ac-CoA from 400 µM at several fixed concentrations of GlcN-1P ranging from 500 to 13 µM.

Further initial velocity pattern experiments were carried out for GlmU acetyltransferase using the CoA-SH release absorbance assay (see Section 2.2.8) by titrating Pro-CoA from 500 µM at several fixed concentrations of GlcN-1P ranging from 6000 to 100 µM.

Acetyltransferase initial velocities were calculated for each condition tested and the data subsequently fitted to both Ping-pong and Ternary complex mechanism models (see Section 2.2.30).

2.2.16 Acetyltransferase substrate specificity

GlmU acetyltransferase Ac-CoA substrate specificity was assessed by utilising the CoA-SH release absorbance assay (see Section 2.2.8) by titrating Ac-CoA analogues from 2 mM in the presence of 2 mM GlcN-1P. Ac-CoA substrate analogues investigated in this study were; AAc-CoA, Pro-CoA, Bu-CoA, Dethio-CoA,, Eth-CoA, IsoBu-CoA, Mlo-CoA, Cro-CoA and Suc-CoA.

GlmU acetyltransferase GlcN-1P substrate specificity was assessed by utilising the CoA-SH release absorbance assay (see Section 2.2.8) by titrating GlcN-1P analogues from 2 mM in the presence of 2 mM Ac-CoA. GlcN-1P substrate

analogues investigated in this study were; GlcN-6P, GalN-1P, GalN-6P, GlcNAc, Glc-1P, GlcN-6P, Gal-1P, GlcN, ManN, ManN-1P, ManNAc, Glc-6P and Man-6P. Acetyltransferase initial velocities were calculated for each condition tested and the data subsequently fitted to the Michaelis-Menten kinetic model (see Section 2.2.30).

2.2.17 Acetyltransferase product and dead-end inhibition

Acetyltransferase inhibition by either the products of the reaction, CoA-SH and GlcNAc-1P, or potential dead-end, substrate analogues, which were selected from the non-substrate acyl-CoAs and sugars assessed in the previous method (see Section 2.2.16). To be able to investigate CoA-SH as a product inhibitor of the acetyltransferase activity, using the previously described thiol detection methodologies, CoA-SH was incubated was an equimolar concentration, typically 5 mM, of *N*-ethylmaleimide (NEM), which was freshly prepared as a concentrated stock in 100 % DMSO, before diluting in assay buffer and then being incubated at 25 °C for 1 hr, prior to use.

Products and potential dead-end analogue inhibition were assessed by utilising the CoA-SH release absorbance assay (see Section 2.2.8) by titrating the test compounds from 10 mM in the presence of 200 µM Ac-CoA or Pro-CoA and GlcN-1P. Acetyltransferase initial velocities were determined for each condition investigated and the data subsequently fitted to a four parameter IC_{50} model (see Section 2.2.30).

2.2.18 Acetyltransferase product and dead-end inhibitor patterns

Acetyltransferase dead-end inhibitor patterns were carried out using either Bu-CoA or Glc-1P. To determine dead-end inhibitor patterns and the inhibition constants (K_{ii} and K_{is}), Bu-CoA or Glc-1P were titrated from 2 mM or 20 mM, respectively, at several fixed concentrations of either Ac-CoA or Glc-1P, within the concentration range of 800 to 10 µM and GlmU activity was monitored using the CoA-SH release absorbance assay (see Section 2.2.8). When varying concentrations of Ac-CoA were investigated, a saturating concentration of 2 mM GlcN-1P was added, alternatively, if GlcN-1P was varied a high concentration of 1 mM Ac-CoA was added. In the case of uncompetitive inhibitor patterns, a variation from the aforementioned methodology is that the matrix titrations of dead-end inhibitor and substrate are

repeated at varying concentrations of the second substrate, which was previously used at a saturating concentration.

Acetyltransferase product inhibitor patterns were carried out using CoA-SH. To determine product inhibitor patterns and the inhibition constants (K_{ii} and K_{is}), CoA-SH was titrated from 1.5 mM at several fixed concentrations of either Ac-CoA or Glc-1P, within the concentration range of 1.5 to 0.15 mM and GlmU acetyltransferase activity was monitored indirectly by coupling through the uridylyltransferase using either the enzyme coupled or PBP Pi accumulation fluorescence assays (see Section 2.2.10). When varying concentrations of Ac-CoA were investigated, a saturating concentration of 2 mM GlcN-1P was added, alternatively, if GlcN-1P was varied a high concentration of 1 mM Ac-CoA was added. To ensure that the uridylyltransferase activity is not limiting, a saturating concentration of 100 µM UTP was added.

For all experiments, acetyltransferase or uridylyltransferase initial velocities were calculated for each condition tested and the data subsequently fitted to one of the three reversible inhibition models (see Section 2.2.30).

2.2.19 Acetyltransferase dead-end inhibitor mutual exclusivity studies

Acetyltransferase dead-end inhibitor mutual exclusivity studies were carried out using Bu-CoA and Glc-1P. Inhibitor mutual exclusivity was assessed utilising the CoA-SH release absorbance assay (see Section 2.2.8) by titrating Bu-CoA from 2 mM at several fixed concentrations of Glc-1P, ranging from 35 to 2.5 mM. Acetyltransferase substrates, 200 µM Ac-CoA and GlcN-1P, were added to GlmU and the dead-end titrations to initiate the experiment. Acetyltransferase initial velocities were determined for each condition investigated and the data subsequently fitted to a Yonetani-Theorell mutual exclusivity model (see Section 2.2.30).

2.2.20 Acetyltransferase pH-rate profiles

To investigate the pH dependence of GlmU acetyltransferase variations of the CoA-SH release absorbance assay (see Section 2.2.8) were utilised in all experiments.

Initially, detection of a titration of CoA-SH from 400 µM by 400 µM DTP was investigated at varying pH from 6 to 10 in a constant ionic strength buffer comprised of 50 mM MES, 50 mM HEPES, 100 mM ethanolamine, 10 mM MgCl2, and 1 mM

CHAPS. This constant ionic strength buffer was used for all pH dependence experiments and the pH adjusted by 0.25 (for GlmU WT) or 0.5 (for GlmU H374A, R344A, K362A and Y377A) increments by additions of either 10 M NaOH or 6 M HCl. GlmU WT and H374A acetyltransferase activity over the pH range from 6 to 10 was investigated by titrating enzyme from a range of concentrations from 400 nM (pH 6 and 10) to 25 nM (pH 7 to 9), with the addition of 200 µM Ac-CoA and GlcN-1P to initiate the reaction, either immediately or after a two hour incubation, at 4 °C, of GlmU in the varying buffer pH conditions. Acetyltransferase initial velocities were determined for each condition investigated and plotted against the concentration of GlmU used.

The pH dependences of $V/K_{GlcN\text{-}1P}$ (k_{cat}/K_m) and V (k_{cat}) were determined by titrating the concentration of GlcN-1P using a range of starting concentrations from 6 to 2 mM (depending on the pH condition) at fixed, saturating concentrations of Ac-CoA between 4 to 2 mM. The pH dependences of $V/K_{Ac\text{-}CoA}$ (k_{cat}/K_m) and V (k_{cat}) was determined by titrating the concentration of Ac-CoA using a range of starting concentrations from 4 to 2 mM (depending on the pH condition) at fixed, saturating concentrations of GlcN-1P between 6 to 2 mM. Acetyltransferase initial velocities were determined for each condition investigated and the data subsequently fitted to the most appropriate pH-rate profile model (see Section 2.2.30).

The pH dependences of GlmU H374A, R344A, K362A and Y377A $V/K_{GlcN\text{-}1P}$ (k_{cat}/K_m) and V (k_{cat}) were determined by titrating the concentration of GlcN-1P from 4 mM at fixed, saturating concentrations of 2 mM Ac-CoA. In parallel, to control for the GlmU H374A studies, GlmU WT $V/K_{GlcN\text{-}1P}$ (k_{cat}/K_m) and V (k_{cat}) were determined by titrating the concentration of GlcN-1P from 4 mM at a fixed, saturating concentration of 2 mM Ac-CoA. The pH dependences of GlmU H374A, R344A, K362A and Y377A $V/K_{Ac\text{-}CoA}$ (k_{cat}/K_m) and V (k_{cat}) were determined by titrating the concentration of Ac-CoA from 2 mM at fixed, saturating concentrations of 4 mM GlcN-1P. Acetyltransferase initial velocities were determined for each condition investigated and the data subsequently fitted to the most appropriate pH-rate profile model (see Section 2.2.30).

2.2.21 Acetyltransferase solvent kinetic isotope effects

Solvent KIEs on *V* and $V/K_{Ac\text{-}CoA}$ were determined using variations of the CoA-SH release absorbance assay (see Section 2.2.8) in water (H_2O)- or 100% deuterium oxide (D_2O)-containing assay buffer, comprising of 50 mM MES, 50 mM HEPES, 100 mM ethanolamine, 10 mM $MgCl_2$, and 1 mM CHAPS, in the presence of saturating concentrations of GlcN-1P at pL 8. Solvent KIEs on *V* and $V/K_{GlcN\text{-}1P}$ were determined in H_2O- or 100% D_2O-containing buffer, comprising of 50 mM MES, 50 mM HEPES, 100 mM ethanolamine, 10 mM $MgCl_2$, and 1 mM CHAPS, in the presence of saturating concentrations of Ac-CoA at pL 8. Viscosity effects on *V*, $V/K_{Ac\text{-}CoA}$ and $V/K_{GlcN\text{-}1P}$ were evaluated by comparing rates obtained in H_2O- or 9% (w/w) glycerol-containing buffer, comprising of 50 mM MES, 50 mM HEPES, 100 mM ethanolamine, 10 mM $MgCl_2$, and 1 mM CHAPS, at pH 8. The use of 9% glycerol mimics the viscosity increase caused by the use of D_2O (η_r = 1.24)[398]. Acetyltransferase initial velocities were determined for each condition investigated and the data subsequently fitted to the most appropriate SKIE model (see Section 2.2.30).

2.2.22 Small molecule compound screening

2.2.22.1 Small molecule compound preparation

Acetyltransferase single-concentration compound preparation for screening was performed by dispensing 1 µl of 1 mM compound solutions in DMSO into clear-bottom, black, 384-well microplates, using an Evo 150 liquid handling instrument (Tecan). The dispense of this volume of test compound gave a final compound concentration of 10 µM in a 100 µl final assay volume. As well as dispensing of test compounds, the same liquid handling instrument was used to dispense 1 µl of 100% dry DMSO to columns 6 and 18 on the 384-well microplates. For concentration-response curves, test compounds were diluted 1 in 3, from 10 mM, over 11-points before transferring 1 µl of every compound concentration to clear-bottom, black, 384-well microplates, all using the Evo 150 liquid handling instrument. The upper concentration in the typical 100 µl acetyltransferase assay was 100 µM.

Uridylyltransferase single-concentration compound preparation for screening was performed by dispensing 200 nl of 1 mM compound solutions in DMSO into small

volume, black, 384-well microplates, using an Echo 555 acoustic liquid handling instrument (Labcyte). The dispense of this volume of test compound gave a final compound concentration of 10 µM in a 20 µl final assay volume. As well as dispensing of test compounds, the same liquid handling instrument was used to dispense 200 nl of 100% dry DMSO to columns 6 and 18 on the 384-well microplates. For concentration-response curves, test compounds were diluted 1 in 3, from 10 mM, over 11-points using the Evo 150 liquid handling instrument before transferring 200 nl of every compound concentration to black, small volume, 384-well microplates, using the Echo acoustic dispenser. The upper concentration in the typical 20 µl uridylyltransferase assay was 100 µM.

The Hewlett-Packard (HP) D300 Digital dispenser (Tecan), was used to prepare bespoke compound concentration ranges for more detailed inhibition studies for both acetyltransferase and uridylyltransferase activities. The D300 dispenses decreasing volumes of compounds in 100% DMSO, from 10 µl to 11 pl directly into the test 384-well microplates and can then be used to backfill with DMSO to ensure a consistent concentration of solvent in every well.

2.2.22.2 Screening process

GlmU acetyltransferase assays were performed by utilising a variation of the CoA-SH release absorbance assay (see Section 2.2.8). Initially, 0.5 ml working enzyme solution, which was prepared in assay buffer and comprised of 5 nM GlmU and 400 µM DTP, was added to all wells containing 1 µl test compound, except column 18, of the test 384-well microplate, using either a Multidrop Micro or TTP Labtech dragonfly discovery (these liquid handlers were used for all subsequent reagent additions). To act as an inhibition or negative control, 0.5 ml of 400 µM DTP in assay buffer, was added to all wells in column 18 of the test microplate. The microplates were then centrifuged for 1 min at 1000 rpm, before incubation at 25 °C, taking care to cover the microplates to minimise evaporation, for 15 mins. To initiate the acetyltransferase reaction, 0.5 ml working substrate solution, which was prepared in assay buffer and comprised of 200 µM GlcN-1P and 100 µM Ac-CoA, was added to all wells of the test 384-well microplate. The microplates were then centrifuged for 10 secs at 1000 rpm and was immediately transferred to a Tecan M1000 PRO or Safire2 monochromator multimode readers and the change in absorbance monitored

at 324 nm ± 5 nm, using 5 flashes per well, at 30 °C. The microplate was monitored for a duration of 30 to 45 mins, using the minimum time interval, determined by the number of wells being read.

Acetyltransferase initial velocities were determined for each condition investigated, then the data were normalised to the positive and negative control columns before subsequently being fitted to a four parameter IC_{50} model (see Section 2.2.30).

To control the quality of the acetyltransferase screening experiments, the mean and standard deviations of the initial velocities of the positive and negative control columns were utilised to determine a Z-factor (see Section 2.2.30)[399].

2.2.23 ^{1}H NMR Spectroscopy and pH titration

NMR measurements were performed at 21.5 °C, on a Bruker Avance 600 MHz equipped with a 5 mm TCI cryoprobe. At each pH value, 1D ^{1}H-noesy[400] and ^{1}H^{13}C-HSQC[401] spectra were acquired on a 40 mM natural abundance GlcN-1P sample in 100 mM KCl, 93% H_2O-7% D_2O solution. The solution contained 2 mM formate, 2 mM imidazole, 2 mM tris and 2 mM piperazine as internal pH indicators, and 0.2 mM DSS for ^{1}H chemical shift internal referencing according to Baryshnikova *et al.*[402].

The pH titration was performed using a cross-titration method where aliquots of the two initial samples of GlcN-1P, adjusted to either pH 5.6 or 11.5, respectively, were reciprocally transferred between samples to achieve intermediate pH values. Spectra were acquired for each point to monitor chemical shifts changes of both GlcN-1P and pH indicators protons. In order to determine pH values at each point of the titration accurately, the chemical shifts of the pH indicators were referred to calibration curves previously acquired at the same temperature, according to the methods described in Oregioni A., *et al.*[403].

H1 and H5 protons of GlcN-1P were preliminary assigned at pH 5.6 to signals at 5.64 and 3.32 ppm, respectively, using a 2D COSY experiment[404], and their chemical shift changes with pH were monitored on the assumption that they, in particular H5, would reflect the amino group ionization state in the measured range of pH.

2.2.24 Differential Scanning Fluorimetry

The ordered, functional structures of proteins are folded to minimise their free energy, which leads to well-packed hydrophobic interiors and hydrophilic exteriors[405].

Differential Scanning Fluorimetry (DSF) also known as fluorescent thermal shift (FTSA), or Thermofluor, is a technique that quantifies the change in thermal denaturation temperature of a protein under varying conditions[406,407]. This technique utilises the intrinsic melting temperature (or T_m) of a protein, which is when both the folded and unfolded states (assuming a two-state model) are equally populated at equilibrium.

SYPRO Orange, a red-shifted fluorophore, binds non-specifically to hydrophobic surfaces, and water strongly quenches its fluorescence[408]. When a protein unfolds, for instance upon the application of a thermal gradient, the exposed hydrophobic surfaces bind the SYPRO Orange, which results in an increase in fluorescence by excluding water (Figure 20).

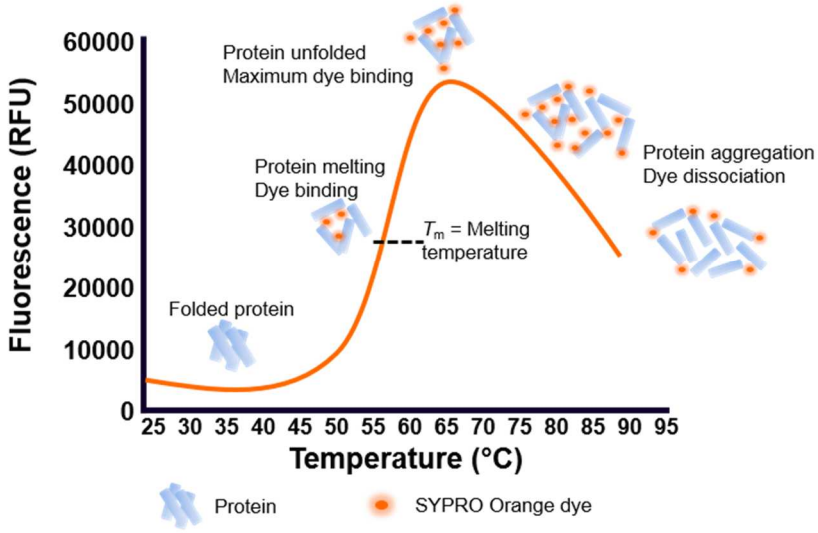


Figure 20. Differential scanning fluorimetry schematic.
Schematic of the basic DSF assay methodology.

The following DSF protocols are variations on the Niesen, *et al.* manuscript published in Nature Protocols[407].

2.2.24.1 Protein T_m determination

Protein unfolding was observed by monitoring the increase in SYPRO Orange (Thermo Fisher Scientific) fluorescence measurements on a Roche LightCycler 480 II RT-PCR machine, using sealed 384-well microplates (white Accuflow Roche 480 plates). SYPRO Orange fluorescence was excited at 483 nm and the emission

measured at 568 nm, taking five reads per 1 °C, as heat was applied to the 384-well microplate to continuously increase the temperature at a ramp rate of 0.11 °C sec^{-1}. A typical experiment was prepared in a buffer comprising of 25 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM CHAPS and GlmU added at a concentration range between 2 to 0.5 µM with a final addition of 10x SYPRO Orange, followed by gentle mixing on an orbital shaker for 5 mins before moving the microplate containing the DSF mix to the PCR machine. SYPRO Orange fluorescence measurements were plotted as a function of temperature for each condition investigated and the data subsequently fitted to Eqn. (19) (see Section 2.2.30) to determine a GlmU T_m.

DSF was used to assess GlmU SDM mutants to assess whether the mutations made have any effect when compared to the WT T_m. To monitor the T_m of GlmU SDM mutants, 1 µM of enzyme was utilised but no other changes were made to the aforementioned protocol.

Control experiments were carried out in the absence of GlmU but in the presence of 10x SYPRO Orange, to monitor any background effects due to either the assay buffer or the PCR microplate.

2.2.24.2 Binding partner and metal ion modulation of T_m

The potential binding partners assessed in these studies included UTP, GlcNAc-1P, CoA-SH, GlcN-1P, GlcN-1P analogues, Ac-CoA and Ac-CoA analogues. DSF studies (see Section 2.2.24.1) were carried out into binding partners ability to modulate GlmU T_m were investigated by titration in the presence of 10x SYPRO Orange and fixed concentrations of enzyme, ranging from 1 to 2 µM. For potential binding partner co-incubation DSF studies, 1 to 2 µM GlmU was incubated with concentrations ranging from 5 to 10 times the K_d concentration of Ac-CoA and CoA-SH for 15 minutes at 4 °C, prior to addition of a titration of the second binding partner and 10x SYPRO Orange.

The divalent metal ions studied in these experiments include $MgCl_2$, $MnCl_2$, $CaCl_2$, $CoCl_2$, $NiCl_2$, $ZnCl_2$ and $Zn(OAc)_2$. DSF studies (see Section 2.2.24.1) were carried out into metal ions ability to modulate GlmU T_m were investigated by titration in the presence of 10x SYPRO Orange and fixed concentrations of enzyme, ranging from 1 to 2 µM. For co-incubation DSF studies with CoA, 1 to 2 µM GlmU was incubated

with concentrations ranging from 5 to 10 times the K_d concentration of CoA-SH for 15 minutes at 4 °C, prior to addition of a titration of metal ion and 10x SYPRO Orange. DSF control studies were carried out utilising either the same binding partner or divalent metal ion titration conditions in the presence of 10x SYPRO Orange and absence of GlmU.

SYPRO Orange fluorescence measurements were plotted as a function of temperature for each condition investigated and the data subsequently fitted to Eqn. (19) model (see Section 2.2.30). The change in GlmU T_m, ΔT_m, were calculated by subtracting the GlmU-only (or GlmU plus substrate binding partner) control T_m from the experimentally determined test conditions GlmU T_m. The GlmU T_m and ΔT_m were then both plotted against the concentration of test metal ion and these were then fitted to a Langmuir Isotherm to determine a K_d.

2.2.24.3 pH modulation of T_m

Differential Scanning Fluorimetry was also utilised to monitor protein stability over a pH range of 6-10, either after a minimal, 15 mins incubation or an extended 2 hr incubation. This study utilised a two-fold concentrate of the following buffer; 50 mM MES, 50mM HEPES, 100 mM Ethanolamine, 10 mM MgCl2, 1 mM CHAPS, which was adjusted to the desired pH using either 6 M HCl or 10 M NaOH. To a white PCR 384-well microplate, a 5 µl addition of 4 µM GlmU was made in a minimal storage buffer [25 mM HEPES (pH 7.4), 100 mM NaCl, 10 mM $MgCl_2$], prior to a 10 µl addition of the pH adjusted assay buffer, which was then gently shaken for 15 minutes. A final addition of 40 % SYPRO Orange, prepared in DI H_2O, was made to all wells, followed by a 10 sec spin at 1000 rpm, prior to either incubating for 2 hrs (and then carrying out the thermal shift step) or carrying out the thermal shift method using the Roche LightCycler 480 II RT-PCR machine immediately. SYPRO Orange fluorescence was excited at 483 nm and the emission measured at 568 nm, taking five reads per 1 °C, as heat was applied to the 384-well microplate to continuously increase the temperature at a ramp rate of 0.11 °C sec^{-1}.

Control experiments were carried out in the absence of GlmU but in the presence of a final assay concertation of 10x SYPRO Orange, to monitor any background effects due to the change in pH of assay buffer.

SYPRO Orange fluorescence measurements were plotted as a function of temperature for each condition investigated and the data subsequently fitted to Eqn. (19) (see Section 2.2.30) to determine a GlmU T_m.

2.2.25 Circular Dichroism spectroscopy

GlmU secondary structure was assessed using Circular Dichroism (CD) spectroscopy using the Chirascan (Applied Photophysics) instrument. Briefly, GlmU was diluted to 0.1 mg/ml in 50 mM HEPES, pH 7.4, and then 320 µl was added to a 1 mm path length cuvette (Hellma, cat. no. 110-1-40) before performing a wavelength scan from 200 nm to 320 nm and measuring CD (mdeg) at increments between 1 to 0.2 nm.

2.2.25.1 CD thermal unfolding

GlmU secondary structure thermal melting (T_m) was investigated by monitoring the CD of the secondary protein structure at a wavelength of 222 nm. GlmU was diluted to 0.1 mg/ml in 50 mM HEPES, pH 7.4, and then 320 µl was added to a 1 mm path length cuvette (Hellma, cat. no. 110-1-40) before performing a temperature ramp of 1 °C sec^{-1} between the range of 8 to 80 °C, while measuring CD (mdeg) at each temperature increment. Data converted to mean residue ellipticity and then Δε, as detailed in the Greenfield Nature Protocols manuscript[409].

2.2.26 NanoDSF

An alternative approach to determine the thermal stability of GlmU prepared in the presence and absence of 10 mM $MgCl_2$, was carried out using NanoDSF, which is a NanoTemper technologies proprietary approach that requires the use of the Prometheus NT.48 instrument, which monitors the melting profiles of the proteins through the increase in intrinsic Tryptophan fluorescence.

For the NanoDSF approach, 30 µL of a 2 µM solution of each protein in assay buffer, either [25 mM HEPES (pH 7.4), 100 mM NaCl, 10 mM $MgCl_2$] or [25 mM HEPES (pH 7.4), 100 mM NaCl], depending on the GlmU preparation, were prepared, and 3 × 10 µl were loaded into nanoDSF grade standard capillaries (NanoTemper Technologies) for triplicate measurements. Thermal unfolding of triplicates was analysed by carrying out a thermal ramp from 25 to 80 °C with a heating rate of 1 °C

min^{-1}. Unfolding transition temperatures (T_m) were automatically determined by the software and represented as mean ± SD.

2.2.27 MicroScale Thermophoresis

Fluorophore labelling of GlmU was performed following the protocol for *N*-hydroxysuccinimide (NHS) coupling of the dye NT647 (NanoTemper Technologies, cat. no. MO-L001) to solvent exposed lysine residues on the surface of the enzyme. Briefly, 3 ml DI H_2O of was added to lyophilised labelling buffer and NT647 was re-suspended by adding 30 µl DMSO to prepare a 435 µM stock. The NT647 fluorophore was subsequently diluted in labelling buffer to prepare a 40 µM stock (taking care to ensure that the DMSO percentage was less than 5%). GlmU was prepared as a 20 µM stock in 100 µl, by diluting in labelling buffer, which was then mixed with 100 µl of the 2 times NT647 fluorophore stock, before incubating for 30 min at 25 °C, taking care to protect the labelling reaction from direct light. After the incubation, 200 µl sample was added to a desalting column, pre-washed with 3 times 3 ml storage buffer [25 mM HEPES (pH 7.4), 100 mM NaCl], before adding 300 µl storage buffer. Elution was then carried out by addition of 600 µl storage buffer to the desalting column before collecting 9 fractions of 1-2 drops per well in a 96-well microplate.

The fluorescence of the collected fractions were measured on the NanoTemper Monolith NT.Automated MicroScale Thermophoresis (MST) instrument, which involved a 1 in 100 dilution of the collected fraction in MST buffer [1% phosphate-buffered saline (PBS), 0.05% P20, 10 mM $MgCl_2$, prepared in DI H_2O], followed by aspiration in standard capillaries and measurement at 20% LED power. The fractions with significant fluorescent signal were pooled and the concentration of both protein and fluorophore were determined, in duplicate, by measuring the absorbance of the sample at 650 nm and 320 nm, a 1 mm pathlength and the extinction coefficient of both the NT647 fluorophore (250000) and GlmU (26025). The measured absorbance values, along with the previously stated parameters, were utilised to determine the concentration of both protein (4.61 µM) and fluorophore (2.54 µM), using Beer-Lambert law (see Section 2.2.30). These data demonstrate that there is approximately one NT647 fluorophore per every two GlmU molecules, which is

recommended by the manufacturer to reduce the chance of any contaminating, unreacted fluorophore.

The interaction between Ac-CoA and NT647-GlmU was established on a Monolith NT.Automated and was subsequently used as a positive control for all following MST studies. In order to characterise this interaction, Ac-CoA serial dilutions from 10 mM, were prepared in H_2O and mixed 3:1 with a solution of 46 nM NT647-GlmU in assay buffer to yield a final volume of 30 µ per dilution, in a black, polystyrene, 384-well microplate (Greiner Bio-One cat. no. 781076). The microplate was then centrifuged for two mins at 2500 rpm before the samples were aspirated into NanoTemper premium capillaries and added to the Monolith where the titration mixes were exposed to light-emitting diode (LED) intensity range between 40 to 60%, at a MST power setting of medium. Stability and reproducibility of the interaction were tested by re-measuring Ac-CoA binding experiments after a 2 hrs incubation time in capillaries at 25 °C. There was no change in fluorescence intensity, protein adsorption, binding amplitude, or K_d value was observed, showing that the interaction was robust, and therefore a suitable positive control for subsequent testing of more potential binding partners.

Test ligands were serially diluted from a top concentration of between 10 to 50 mM before being mixed at either 1:1 or 3:1 with 46 nM NT647-GlmU in assay buffer to yield a final volume of 30 µ per dilution, in a black, polystyrene, 384-well microplate (Greiner Bio-One cat. no. 781076). The test samples were then treated as the previously described Ac-CoA control, including the capillaries used and the Monolith settings.

To determine K_d values for potential binding partners concentration-dependent changes in normalised fluorescence (F_{norm}) of NT647-GlmU after 5 secs of thermophoresis based on the law of mass action using the NT.Affinity Analysis software.

2.2.28 Minimal inhibitory concentration determination

The direct measurement of the minimum inhibitory concentration (MIC) for each tested compound on *Mycobacterium tuberculosis* in culture, were performed in flat-bottom, polystyrene, 96-well microplates. Test compound twofold dilutions were carried out from 50 mM in 100% DMSO and then 5 µl was added to 95 µl Middlebrook

7H9 medium (rows A–H, columns 1–10 of the microplate layout). Isoniazid was used as a positive control and eight twofold dilutions, starting at 160 µg ml^{-1} were prepared and 5 µl of the control dilution was added to 95 µl Middlebrook 7H9 medium (column 11, rows A–H). A positive control of 5 µl of 100% DMSO was added to 95 µl Middlebrook 7H9 medium to column 12. The inoculum was standardised to ~1×10^7 CFU ml^{-1} and diluted 1 in 100 in Middlebrook 7H9 broth (Middlebrook ADC enrichment, a dehydrated culture medium which supports growth of mycobacterial species, available from Becton–Dickinson, cat. no. 211887), to produce the final inoculum of H37Rv strain (ATCC25618). Subsequently, 100 µl of the inoculum was added to all the wells on the test microplate except G12 and H12 wells, which were blank controls. All plates were placed in a sealed box to prevent drying out of the peripheral microplate wells and were incubated at 37 °C without shaking for six days. A detection solution was prepared by dissolving one tablet of resazurin (VWR International, *Resazurin Tablets for Milk Testing*, cat. no. 330884Y′) in 30 ml sterile PBS and then 25 µl was added to each well on the microplate. After 48 hrs incubation at 37 °C, fluorescence was measured by excitation at 530 nm and monitoring emission at 590 nm using a SPECTRAmax M5 (Molecular Devices) and the data was then aligned to the test compound concentrations and fitted to determine the minimum inhibitory concentration (MIC) value. The MIC is defined as the lowest concentration of a test compound that prevents visible growth of bacterium.

In addition to the direct measurement of MIC on *Mycobacterium tuberculosis*, using the aforementioned fluorescence intensity measurement, sometimes referred to as the extracellular MIC assay, an intracellular assay was carried out. The intracellular assay monitors the MIC of *M. tuberculosis* H37Rv-infected human THP-1 cells, which is a monocytic cell line derived from an acute monocytic leukaemia patient.

Briefly, *M. tuberculosis* H37Rv containing the *Photinus pyralis* luciferase gene (Hygromycin resistant plasmid) was grown in 7H9 media supplemented with 10% albumin-dextrose-catalase (ADC) and 0.05% Tyloxapol until an OD_{600} of 0.5 – 0.8 was achieved. The *M. tuberculosis* culture was then divided into centrifuge tubes and then pelleted by centrifugation at 2860 *x g* for 10 mins. To disperse the bacterial pellet; 10, 4 mm, glass beads were added to each centrifuge tube before shaking for 60 secs. After adding fresh RPMI (Roswell Park Memorial Institute) media, the culture was incubated for 5 mins, before collecting 5 ml of the supernatant from each centrifuge tube, pooling and centrifugation for 5 mins at 402 *x g*. The subsequent

dispersed bacterial suspension was diluted in RPMI media, supplemented in 0.05% Tyloxapol, before the volume required for multiplicity of infection (MOI) of 1 was calculated using the following conversion; OD_{600} 0.1 = 1 x 10^7 colony-forming unit (CFU)/ml. THP-1 cells were maintained in complete RPMI 1640 media (RPMI 1640 HEPES modification, 2 mM L-glutamine, 1 mM sodium pyruvate, 10% foetal bovine serum) and incubated at 37 °C with 5% CO_2. THP-1 phagocytes (2 x 10^5 cell/ml) were incubated for 4 hrs in a roller bottle with a MOI of 1 in RPMI, supplemented with 20 nM Phorbol-12-Myristate-13-Acetate (PMA). The remaining extracellular bacteria were discarded by washing 5 times in complete RPMI, after 5 cycles of gentle centrifugation for 5 mins at 402 *x g*. To a white 384-well microplate, 50 µl of infected THP-1 cells (10000 cells/well) were added to 250 nl of test compounds, prepared in 100% DMSO. Subsequently, the test microplates were incubated for 5 days in an incubator set to 37 °C in the presence of 5% CO_2, before adding 25 µl of reconstituted Bright-Glo™ Luciferase Assay System reagent (Promega) to each well. The test microplates were incubated at 25 °C for 30 mins, prior to measurement of the luminescence signal using a Perkin Elmer Envision microplate reader, equipped with an Ultra-sensitive luminescence detector and a 384-well microplate aperture.

2.2.29 Crystallisation, data collection and refinement of GlmU bound to Glc-1P and Ac-CoA

To grow GlmU crystals, the protein solution in storage buffer [25 mM HEPES (pH 7.4), 100 mM NaCl, 10 mM $MgCl_2$] was concentrated to 11 mg/ml using Amicon Ultra filters and the centrifugation method detailed in the “General methods and equipment” section. The apo-GlmU protein was crystallised at 20 °C using the sitting-drop vapour diffusion method[410], where sitting drops of 1 µl consisted of a 1:1 (v/v) mixture of protein and a well solution containing 0.05 M ADA (pH 6.8), 5.7 % Polyethylene glycol (PEG) 550 MME and 32.1 % PEG 200. Crystals appeared after 12 hrs and reached their maximum size after 2 days (100 µm x 100 µm x 40 µm). Crystals were then soaked for 5 min in a solution containing 10 mM Glc-1P, 10 mM Ac-CoA, 0.05 M ADA (pH 6.8), 8 % PEG 550 MME and 35 % PEG 200 Crystals were directly flash-frozen in liquid nitrogen, and X-ray data sets were collected at 100 K at the I03 beamline of the Diamond Light Source Synchrotron (Oxford, UK, mx13775-39).

Data collection and refinement statistics are summarized in Table 17. The data set was indexed, scaled and merged with xia2[411]. Molecular replacement was achieved by using the atomic coordinates of *Mycobacterium tuberculosis* GlmU from pdb 4G87[246] in PHASER[412]. Refinement was carried out by using Phenix[413], model building was carried out in COOT[414], model validation used PROCHECK[415] and figures were prepared using the graphics program PYMOL[416]. The asymmetric unit contains one chain of GlmU. The difference electron density map covering GlmU shows unambiguous density for Glc-1P and Ac-CoA.

2.2.30 Data analysis

All data handling was carried out in Microsoft Excel 2016. A linear model (y=((A*x) + B)) was used to determine intercept (A) and gradient (B) of kinetic data and was carried out using the Excel addin IDBS XLFit 5.5.0.5. All initial rate data were fitted using either SigmaPlot 12.5 or Grafit 7.0.2. Errors were propagated as described in Skoog and West for indeterminate errors[417]. Individual saturation curves were fitted to (1),

(1)

$$\boldsymbol{v = VA/(A + K)}$$

where *V* is the maximal velocity, *A* is the substrate concentration, and *K* is the Michaelis constant for the substrate (K_m). Saturation curve showing a non-zero *Y* intercept was fitted to (2),

(2)

$$v = v_0 + [VA/(A + K)]$$

where v_0 is the velocity in the absence of activator, *V* is the maximal velocity, *A* is the concentration of activator, and *K* is the concentration of activator that gives half-maximal activation (K_{act}). Individual saturating curves showing linear substrate inhibition were fitted to (3)

(3)

$$v = VA/(K + A\ (A^2/K_i\ )\ )$$

where K_i is the apparent inhibition constant for substrate *A*. Data showing an intersecting initial velocity pattern on double-reciprocal plots were fitted to (4),

(4)

$$v = VAB/(K_{ia}K_B + K_AB + K_BA + AB)$$

where *A* and *B* are the concentrations of the substrates and K_A and K_B are the Michaelis constants. Inhibition data showing linear, competitive, non-competitive or uncompetitive patterns in double-reciprocal plots were fitted to (5), (6) and (7), respectively,

(5)

$$v = VA/[K(1 + I/K_{is}) + A]$$

(6)

$$v = VA/[K(1 + I/K_{is}) + A(1 + I/K_{ii})]$$

(7)

$$v = VA/[K + A(1 + I/K_{ii})]$$

where *I* is the inhibitor concentration and K_{is} and K_{ii} are the slope and intercept inhibition constants, respectively. Inhibition data were fitted to (8),

(8)

$$v = v_0/[1 + (I/IC_{50})^{nH} + D]$$

where *v* is the rate in the presence of the inhibitor at concentration *I*, v_0 is the rate without the inhibitor, IC_{50} is the concentration of the inhibitor that gives 50% inhibition, n_H is the Hill coefficient and *D* is the assay background. pH profile data were fitted to (9) for one basic ionizable group, (10) for two nonresolvable acidic ionizable groups, (11) for two acidic nonresolvable and two basic nonresolvable ionizable groups, (12), for two acidic nonresolvable and one basic ionizable group, (13) for one acidic ionizable group and (14) for one acidic and two basic nonresolvable ionizable groups,

(9)

$$v = C/(1 + H/K_b)$$

(10)

$$v = C/(1 + H^2/K_a^2)$$

(11)

$$v = C/[1 + (H^2/K_a^2) + (K_b^2/H^2)]$$

(12)

$$v = C/[1 + (H^2/K_a) + K_b/H]$$

(13)

$$v = C/(1 + H/K_a)$$

(14)

$$v = C/[1 + (H/K_a) + (K_b^2/H^2)]$$

where *C* is the pH-independent plateau value, *H* is the hydrogen ion concentration, and K_a and K_b are the respective acidic and basic pK_a constants for the ionizable groups. Solvent kinetic isotope effects were fitted to (15, (16, and (17, for isotope effects on *V* only, *V/K* only, or both *V* and *V/K*, respectively,

(15)

$$\nu = VA/[K + A(1 + F_i E_V)]$$

(16)

$$\nu = VA/[K(1 + F_i E_V/K) + A]$$

(17)

$$\nu = VA/[K(1 + F_i E_V/\mathrm{K}) + A(1 + F_i E_V)]$$

where *F*i is the fraction of the isotopic label and E_V and $E_{V/K}$ are the isotope effects minus one on *V* and *V/K*, respectively. Linear V_{max} proton inventory data were fitted to (18),

(18)

$$V_n = V_n[1 - n + n(k_D/k_H)]$$

where n is the atom fraction of deuterium, V_n is the velocity in the solvent with the atom fraction of deuterium n, V_0 is the velocity in H_2O, and k_D/k_H is the isotope effect. Differential scanning fluorimetry thermal melting experiments were fitted to (19),

(19)

$$y = \mathrm{LL} + (\mathrm{UL} - \mathrm{LL})/[1 + \exp(T_m - X/E)]$$

where LL and UL are the values of minimum and maximum intensities, respectively, *X* is the temperature, *E* is the slope of the curve and T_m is the melting point. Plots of change in T_m (ΔT_m) versus ligand concentration were fitted to (20)

(20)

$$\Delta T_m = \Delta T_{m,max} A/(K_d + A) + D$$

where *A* is the ligand being titrated, K_d is its apparent dissociation constant, *D* is the assay background, and ΔT_{mmax} is the maximum possible T_m change. pH dependence of ^{1}H GlcN-1P chemical shifts were fitted to (21),

(21)

$$\delta_{peak} = \delta_{HA} + (\Delta\delta/\ 1 + 10^{\mathrm{nH(pH\text{-}pKa)}})$$

where δ_{peak} is the peak chemical shift, is δ_{HA} is the chemical shift of the protonated form and $\Delta\delta$ is the difference between δ_{HA} and the shift of the deprotonated form. Mutual exclusivity data were fitted to (22),

(22)

$$v = V/[(1 + (I/K_i) + (J/K_j) + IJ)/(\alpha K_i K_j)]$$

where *V* is the maximal velocity, *I* and *J* are the inhibitor concentrations, K_i and K_j are the dissociation constants for the inhibitors, and *α* is an interaction term that defines the effect of the binding of one inhibitor on the affinity of the second inhibitor. In order to normalise the assay specific response in compound screening experiments, the means of the specific microplate positive and negative controls, were utilised to convert the response to a percentage inhibition (%I) by applying (23),

(23)

$$\%I = (\mu_1 - X)/(\mu_1 - \mu_2)) * 100$$

where *X* is the assay specific response, μ_1 is the mean of the negative, or uninhibited control and μ_2 is the mean of the positive, or inhibited control. Compound screening assay quality was monitored by using two control populations on every microplate tested to calculate a "Z'-factor", which is screening window coefficient and is defined by (24),

(24)

$$Z' = 1 - (((3 * \sigma_1) + (3 * \sigma_2))/(\mu_1 - \mu_2))$$

where σ_1 and σ_2 are the standard deviations of the positive and negative controls, μ_1 is the mean of the negative, or uninhibited control and μ_2 is the mean of the positive, or inhibited control. The MST signal detects the binding of a test compound by quantification of the change in the normalised fluorescence (i.e., the amplitude of the MST signal). The fraction of labelled target bound to their test compound, x, enables the change in fluorescent signal that depends on the concentration of target is determined by (25)

(25)

$$F_{\text{norm}} = (1 - x)F_{\text{norm}}(unbound) + x\,F_{\text{norm}}(bound)$$

where, F_{norm} (unbound) is the normalised fluorescence of unbound labelled molecules, and F_{norm} (bound) is the normalised fluorescence of complexes (*i.e.*, of labelled molecules bound to their targets in saturation). The unbound state acts as a reference state. The dissociation constant, K_d, was determined for ligands binding to GlmU were fitted to (26),

(26)

$$AB = (A_0 + B)/(K_d + B)$$

where AB is the concentration of the bound complex, A_0 is the total amount of one binding molecule added, B is the free concentration of the second binding molecule and K_d is the dissociation complex. The half maximal effective concentration, EC_{50}, was determined for a ligand-mediated effect on GlmU are was fitted to (27),

(27)

$$v = v_0/[1 + (I/EC_{50})^{n\mathrm{H}} + D]$$

where *v* is the rate in the presence of the activator at concentration *I*, v_0 is the rate without the activator, EC_{50} is the concentration of the inhibitor that gives 50% activation, n_H is the Hill coefficient and *D* is the assay background. Whole-cell antimicrobial activity was determined by broth microdilution using the Clinical and Laboratory Standards Institute (CLSI) recommended procedure, Document M7-A7, "Methods for Dilution Susceptibility Tests for Bacteria that Grow Aerobically". The minimum inhibitor concentration (MIC) lowest concentration of compound required to produce a >80% decrease in observed fluorescence or luminescence. Kinetic absorbance data, measured as an OD, were converted to concentration of product by used of the Beer-Lambert equation (28),

(28)

$$A = \varepsilon lc$$

where *A* is the absorbance, *ε* is the molar absorption coefficient, *l* is the pathlength of the solution (cm) and *c* is the concentration of the solution (mol cm^{-3}). Pearson correlation coefficient equation was used to exclude non-linear kinetic data (29),

(29)

$$r = \Sigma((x - \bar{x})(y - \bar{y}))/\sqrt{\Sigma(x - \bar{x})^2\,\Sigma(y - \bar{y})^2}$$

where *x* is the time (secs) and *y* is the concentration of product formed ([CoA-SH] (μM) or [Pi] (μM)). A linear regression equation was used to determine the concentration of product formed per unit time (usually secs) (30),

(30)

$$y = bx + a$$

where *x* is the time (secs), *y* is the concentration of product formed ([CoA-SH] (μM) or [Pi] (μM)), *b* is the gradient and *a* is the *y*-axis intercept.

Chapter 3. Characterisation of the kinetic mechanism of GlmU catalysed acetyl transfer

To better understand the catalytic activities of GlmU, studies were undertaken to thoroughly, kinetically characterise the acetyltransferase activity. Similar studies have been carried out previously both on the *M. tuberculosis* enzyme and related bacterial orthologues, however, most of these used discontinuous approaches and were not detailed enough for the purposes of the future work, including novel inhibitor identification and confirmation of GlmU essentiality *in vitro*[215,245]. A detailed examination of the preparation of GlmU and subsequent identity of the protein, were carried out to ensure that the optimal conditions were utilised. Likewise, the proposed assay approaches to monitor acetyltransferase activity were thoroughly compared to ensure that the most appropriate methodology was taken forward for further studies. Divalent metal ions have previously been observed at the GlmU acetyltransferase trimer interface and the effect of these on both protein stability and acetyltransferase were studied[220]. To understand the order of substrate binding and product release, acetyltransferase kinetic mechanism was elucidated using the following studies; initial velocity patterns, detailed examination of the specificity of the substrates and finally dead-end analogue and product inhibitor experiments. To complement and expand on the kinetic examination of the GlmU acetyltransferase activity, substrate and product binding studies were carried out using DSF, as well as X-ray crystallography to understand the structure and conformation of the protein-substrate catalytic complex. The structure of free GlmU, along with several ligand-bound forms have previously been solved[245], however these studies have captured the catalytic complex pre-acetyl-transfer, which would provide a better representation of the correct orientation of both the active site residues and substrates, prior to catalysis. All equations used in the analysis of data in this chapter can be found in the Materials and Methods chapter, section 2.2.30.

3.1 Protein preparation and quantification

Protein expression was carried out using both the *N*- and *C*-terminal hexahistidine, TEV cleavage site constructs to generate two 120 grams of cell pellet. Purification of the *C*-terminal hexahistidine construct was carried out first and SDS-PAGE gel

electrophoresis analysis of the final product highlighted the presence of two significant species, identified by the bands in Figure 21.

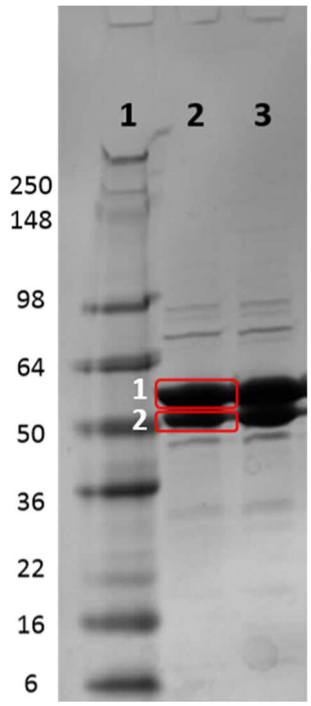


Figure 21. SDS-PAGE gel electrophoresis image of GlmU-TEV-His_6

SDS-PAGE analysis of the final purification product of GlmU-TEV-His_6. The two bands, highlighted by two red squares (lane 2), are approximately the expected molecular weight (MWt) for this C-terminal GlmU construct.

After GlmU-TEV-His_6 final purification, intact mass was measured by ESI-LC-MS, which yielded an average molecular weight (MWt) of 53.4 kDa. The expected MWt for this construct was 53.41 kDa and the measured mass is consistent with the loss of the *N*-terminal methionine. Protein concentration was measured as 13.2 mg/ml and a total of 132 mgs was purified from 30 grams of starting cell pellet. Further analysis of the species of GlmU separated by SDS-PAGE gel electrophoresis was carried out by Peptide Mass-Fingerprinting (PMF), which identified both as GlmU with an average of 27 % peptide coverage after Trypsin cleavage. Comparison of the PMF data for bands identified in Figure 21 as 1 and 2, highlights the absence of peptides covering the C-terminal of band 2. These data indicate that there is significant cleavage of GlmU-TEV-His_6, leading to a heterogenous population that cannot be separated by the chromatography method utilised for protein purification. Further protein preparation was carried out using the His_6-TEV-GlmU construct and purified protein was analysed by SDS-PAGE gel electrophoresis (Figure 22), which

highlights a clear difference, in terms of number of significant bands per lane, between *N*- and *C*-terminal hexahistidine, TEV cleavage site constructs.

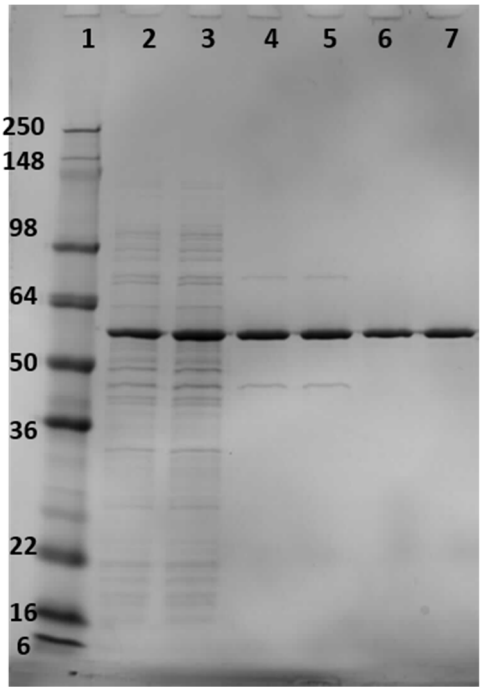


Figure 22. SDS-PAGE gel electrophoresis analysis of His_6-TEV-GlmU.

SDS-PAGE analysis of the final purification products of His_6-TEV-GlmU. The gel layout is as follows; lane 1 is the gel MWt marker, lanes 2 and 3 are the SEC load, lanes 4 and 5 are the first SEC pool and lanes 6 and 7 are the second SEC pool.

The protein accurate mass was determined by ESI LC-MS as 53.4 kDa, which is consistent with *N*-terminal methionine cleavage. Protein concentration was measured as 5.3 mg/ml and a total of 89 mgs was purified from 30 g of starting cell pellet. Protein identity was confirmed as GlmU using PMF analysis, which identified 22 peptides after trypsin cleavage, from the single bands in gel lanes 6 and 7 (Figure 22), equating to 65% coverage of the expected amino acid sequence. These data confirm the homogeneity, sequence and mass of the protein purified from *E. coli* cell pellet as His_6-TEV-GlmU.

To analyse the oligomeric state of GlmU gel electrophoresis using non-reducing SDS-PAGE was performed (removal of DTT reducing agent from sample preparation buffer) and native-PAGE approaches (Figure 23 A and B).

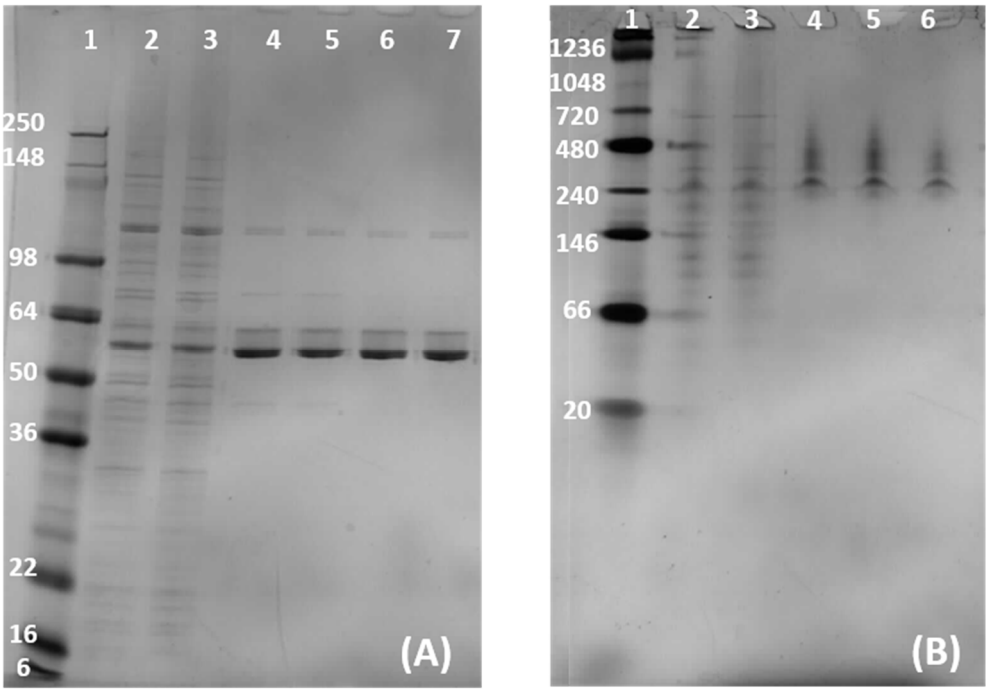


Figure 23. Further gel electrophoresis analysis of His_6-TEV-GlmU.

Gel electrophoresis analysis, (A) Non-reducing SDS-PAGE and (B) Native-PAGE, of the final purification products of His_6-TEV-GlmU. (A) The non-reducing SDS-PAGE gel layout is as follows; lane 1 is the gel MWt marker, lanes 2 and 3 are the SEC load, lanes 4 and 5 are the first SEC pool and lanes 6 and 7 are the second SEC pool. (B) The Native-PAGE gel layout is as follows; lane 1 is the gel MWt marker, lanes 2 and 3 are the SEC load, lanes 4 and 5 are the first SEC pool and lane 6 and 7 is the second SEC pool.

The non-reducing SDS-PAGE analysis of His_6-TEV-GlmU (Figure 23 A) demonstrates that there are two bands at ~ 53.4 kDa per GlmU sample lane, one containing the majority of sample and another minor band at a higher MWt. These data indicate that any quaternary GlmU structure is not significantly stabilised by disulphide bridges, as the observed bands are at the expected position between 50 - 64 kDa on the gel. The appearance of a minor band at ~ 60 kDa, when compared to the reducing SDS-PAGE analysis, may indicate the presence of minor disulphide bridges that perturb the migration of the denatured sample down the gel. There are no significant higher MWt bands present, which is indicative of the absence of any covalently linked complexes. The native-PAGE analysis of His_6-TEV-GlmU (Figure 23 B) reveals the presence of a significant band at ~ 240-300 kDa in the sample lanes containing the GlmU samples. Previous studies have proposed that GlmU is a homotrimer in solution, primarily based on data from structural biology studies

using X-ray crystallography. The data presented in Figure 23 confirm that GlmU exists in solution as a multimeric complex and given the monomer is 53.4 kDa, then these findings are indicative of the existence a hexamer. The presence of the sample added to the Native-PAGE gel in the GlmU preparation lanes has migrated into the gel, in contrast to the SEC load lanes, where a significant amount of sample has not migrated into the gel and can still be observed in the loading wells.

The final pooled purification product of His_6-TEV-GlmU was analysed by CD spectroscopy to confirm the presence of secondary polypeptide structure (Figure 24).

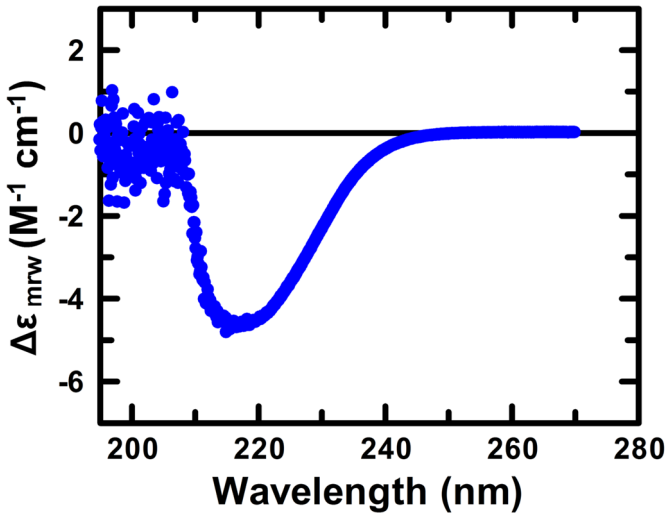


Figure 24. Far-UV CD spectra of His_6-TEV-GlmU.

His_6-TEV-GlmU protein secondary structure spectrum was determined by far-UV CD spectroscopy, using a wavelength scan from 190 – 270 nm. These data are a mean of three discrete replicates that have been background subtracted using a buffer blank measurement.

The His_6-TEV-GlmU far-UV CD spectrum determined is indicative of a folded protein secondary structure that has a significant proportion of antiparallel β-sheet[409] and more specifically, parallel β-helix[418]. These secondary structure data describe the composition of the published X-ray crystallography structures of GlmU, which demonstrate that the preparation of *M. tuberculosis* enzyme for this study is likely optimally folded. The significant noise observed at wavelengths lower than 210 nm was due to the presence of a relatively high concentration of Cl^- ions in the protein storage buffer[419]. The CD spectrum of His_6-TEV-GlmU was used to identify a suitable wavelength to monitor during a thermal unfolding experiment to characterise the thermal stability of GlmU and determine a T_m (Figure 25).

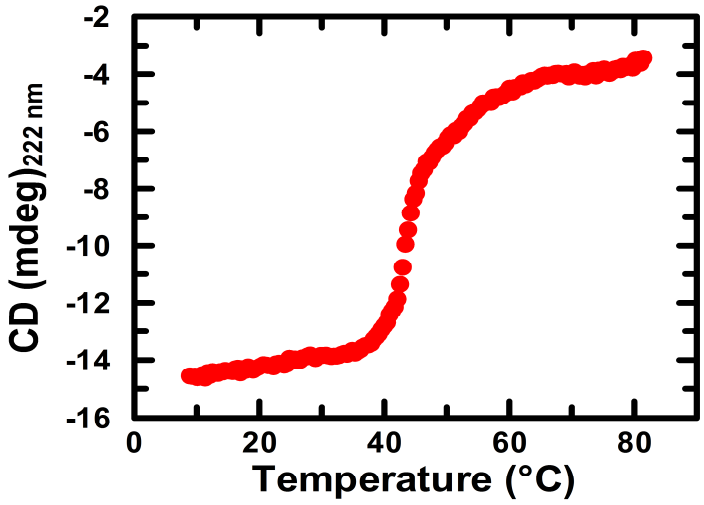


Figure 25. His_6-TEV-GlmU Far-UV CD thermal unfolding.

CD signal of the secondary structure of His_6-TEV-GlmU protein at 222 nm was used to monitor protein structural integrity at increasing temperature from 4 to 80 °C. The data have been fitted, where appropriate, to Eqn. (19).

The secondary structure of His_6-TEV-GlmU was further characterised by investigating the change in protein CD at 222 nm over a temperature gradient to determine the melting point, T_m. The thermal unfolding data in Figure 25. was fitted to Eqn. (19) to determine a His_6-TEV-GlmU T_m of 45.1 ± 0.1 °C, which further demonstrates that the pooled preparation of His_6-TEV-GlmU is folded and has significant secondary structure.

3.2 Acetyltransferase absorbance assay configuration and non-linear initial velocities

The acetyltransferase activity of GlmU was monitored using free thiol detection assays to detect the accumulation of the product CoA-SH in real time. The two detection methodologies initially employed to measure the free thiol of the acetyltransferase product, CoA-SH, were both chromogenic reagents; the classically utilised DTNB (also known as Ellman's reagent) and DTP (also known as Aldrithiol-4). The two detection reagents have highly oxidising disulphide bonds that are stoichiometrically reduced by free thiols in an exchange reaction, forming mixed disulphides and releasing one molecule of either NTB^{2-} or 4-TP, respectively. NTB^{2-} and 4-TP are good leaving groups and in all cases one of either NTB^{2-} or 4-TP are released for every thiol oxidised in the presence of either DTNB or DTP,

respectively[420,421]. The formation of photometrically detectable NTB^{2-} or 4-TP, leads to an increase in optical density at either 412 or 324 nm, respectively.

To initially compare the application of DTNB and DTP to detect thiols, CoA-SH was titrated in assay buffer at pH 7.5, prior to addition of the detection reagent (Figure 26).

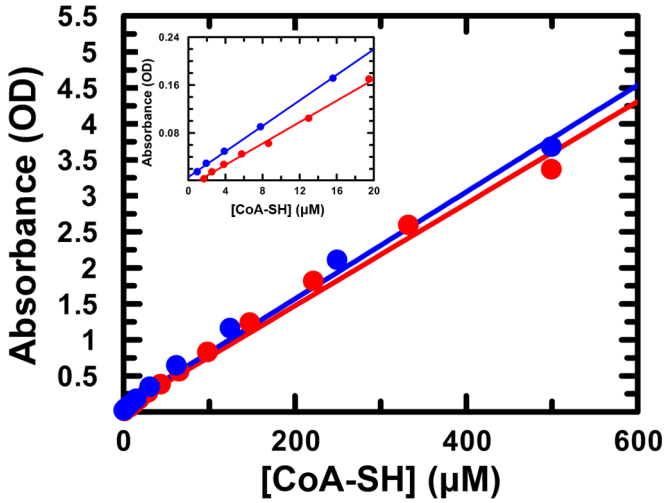


Figure 26. Thiol detection of CoA-SH titrations.

CoA-SH was titrated in assay buffer at pH 7.5 and subsequently detected by additions of 500 μM of either DTNB (●) or DTP (●). Absorbance was monitored at an optical density (OD) of either 324 nm (DTP) or 412 nm (DTNB). Each of these data have been fitted to a linear regression; DTNB (—, correlation coefficient = 0.99) or DTP (—, correlation coefficient = 0.99). Inset figure, CoA-SH concentration range from 1 to 20 μM.

The use of DTNB and DTP to detect the free thiol of CoA-SH were equivalent at pH 7.5, with a minimum level of detection of 1 μM that was linear up to a concentration of 500 μM thiol (Figure 26 Inset).

As a final control, before proceeding with the use of either DTNB or DTP to monitor acetyltransferase activity, GlmU was titrated in assay buffer at pH 7.5, in the absence of substrates but in the presence of thiol detection reagents. These data, utilising a range of GlmU concentrations between 4 nM and 5 μM, showed that no increase in absorbance was observed above background. These findings demonstrate that any background absorbance is not due to the addition of GlmU and subsequent reaction of surface accessible thiols with the detection reagents.

To investigate whether any surface accessible thiols are essential for acetyltransferase activity, a concentrated stock of GlmU was incubated with an

equimolar concentration of maleimide, *N*-ethylmaleimide (NEM), for 30 mins, at 25 °C, prior to titration in the presence of 200 µM of both Ac-CoA and GlcN-1P and the use of DTP to monitor activity by the change in absorbance. To control this experiment, an additional titration of non-NEM-incubated GlmU was carried out in the presence of the acetyltransferase substrates and activity monitored using the DTP absorbance assay (Figure 27).

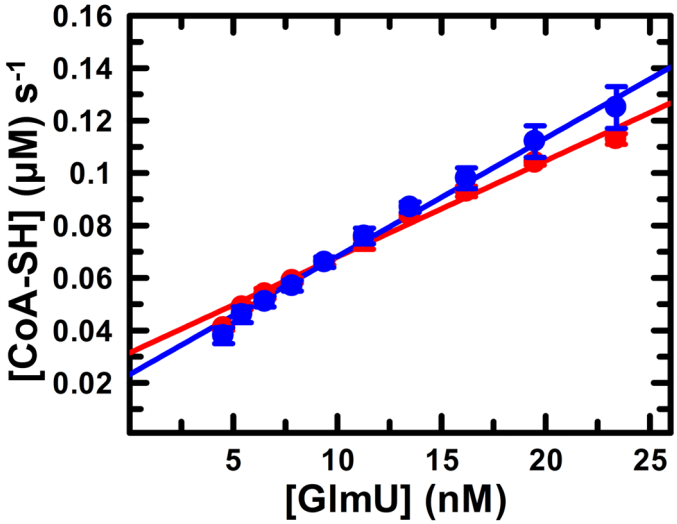


Figure 27. Validation of thiol detection assays to monitor GlmU acetyltransferase activity.

NEM-incubated (●) and non-modified GlmU (●) were titrated in assay buffer at pH 7.5 in the presence of DTP, before addition of 200 µM acetyltransferase substrates to determine if any solvent accessible thiols are catalytically essential. These data are a mean of six discrete replicates and the solid lines are fits of the experimental data to a linear regression model.

The control experiments in Figure 27 demonstrated that the pre-incubation of GlmU with NEM and subsequent detection of CoA-SH formation using DTP, doesn't prevent acetyltransferase and demonstrates that the solvent exposed thiol groups are not involved in catalysis. Furthermore, there was no significant difference between the acetyltransferase activities of NEM-treated and unmodified GlmU, which further confirms that modification of solvent exposed thiols on the enzyme have no effect on acetyl transfer. The quaternary structure of GlmU is not formed through disulphide bridge formation and the exposure of the enzyme to thiol modification reagents is not likely to disrupt the trimer. Each of the three acetyltransferase active sites involve contributions from all monomers in the trimer[215]. If the trimer were to be disrupted by exposure to high concentrations of thiol

modification reagents, then GlmU acetyltransferase activity and CoA-SH formation would not be possible. To further investigate and evaluate the thiol detection methodologies, DTP and DTNB were used to monitor the acetyltransferase activity of GlmU by kinetically detecting the formation of CoA-SH in assay buffer at pH 7.5 (Figure 28).

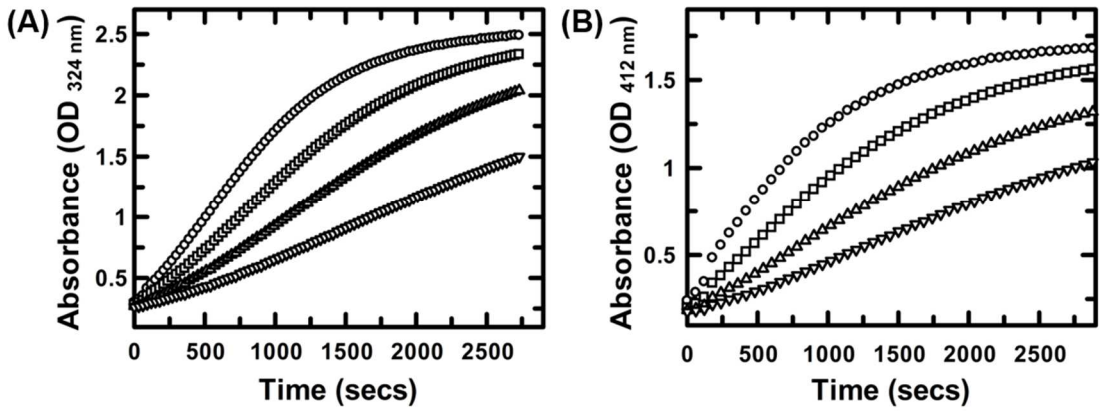


Figure 28. Acetyltransferase activity of a titration of GlmU.

GlmU was titrated in the presence of 200 µM of both Ac-CoA and GlcN-1P in assay buffer at pH 7.5. Accumulation of the product CoA-SH was monitored by absorbance using 500 µM of either (A) DTP or (B) DTNB. The concentrations of GlmU were; (○) 12 nM, (□) 6 nM, (▲) 3 nM and (▼) 1.5 nM.

There was no difference between the use of either DTP or DTNB to monitor GlmU acetyltransferase activity kinetically, however, with both detection reagents it was apparent there was an initial non-linear phase, up to 500 secs. To investigate this lag phase, which could be indicative of a slow-dissociation of a competitive or a slow transition to an active enzyme conformation, several factors were considered, including inhibition by kanamycin (derived from expression phase) and temperature effects[422,423]. In addition, to try and isolate any unknown, tightly bound inhibitors, concentrated GlmU was denatured at 95 °C, prior to removal of aggregate by centrifugation, before titrating the supernatant in the presence of 10 nM GlmU and incubating for 60 mins at 25 °C. Acetyltransferase activity was then monitored using the DTP absorbance assay, using a buffer at pH 7.5, containing 10 mM $MgCl_2$ in the presence of fixed concentrations of substrates. None of the aforementioned approaches resulted in the identification of the cause of the non-linear acetyltransferase kinetics. Incubation of concentrated GlmU with 10 mM $MgCl_2$ at 4

°C, prior to assay, was identified as a critical requirement for optimal, linear acetyltransferase activity (Figure 29).

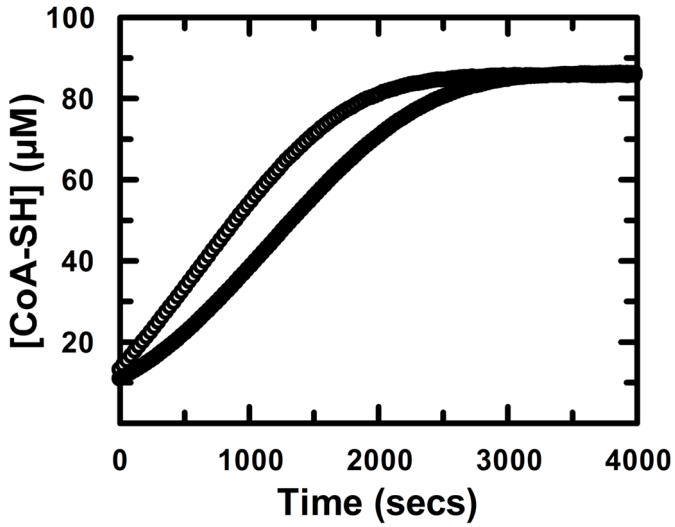


Figure 29. Acetyltransferase activity of GlmU in the presence of 10 mM $MgCl_2$.

Progress curves of the GlmU acetyltransferase reaction using DTP to monitor formation of CoA-SH. The change in absorbance of DTP at OD_{324nm} was monitored every 10 secs and then converted to concentration of CoA-SH formed using Beer-Lambert law (see Materials and Methods, "Data analysis" section). The data for a high concentration of GlmU incubated in the presence or absence of 10 mM $MgCl_2$, then used in the assay at 10 nM, are represented by either ○ or ●, respectively.

The non-linear kinetics of GlmU acetyltransferase, evident in both Figure 28 and Figure 29, appear to be eradicated by an initial extended incubation, for periods in excess of 20 hours, of protein with 10 mM $MgCl_2$ at 4 °C, prior to assay. To further investigate the apparent dependency of GlmU for 10 mM $MgCl_2$ for optimal acetyltransferase activity, an incubation of a high concentration of protein with either 10 mM EDTA or $MgCl_2$ for 12 hrs at 4 °C was carried out, prior to monitoring the acetyltransferase activity (Figure 30).

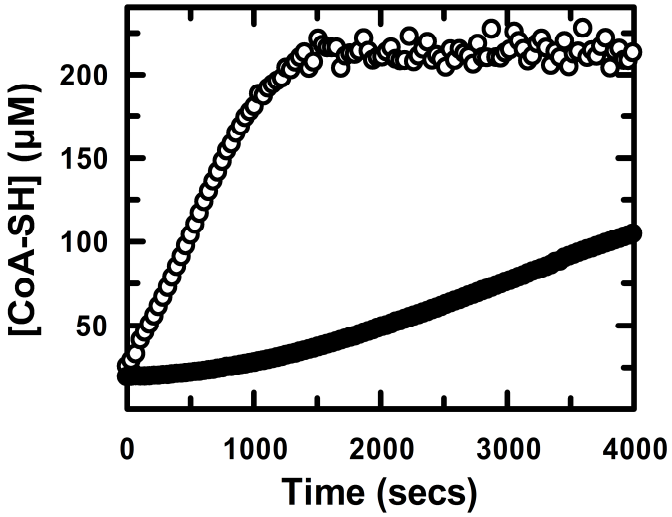


Figure 30. Acetyltransferase activity of GlmU in the presence of 10 mM EDTA.

Progress curves of the GlmU acetyltransferase reaction using DTP to monitor formation of CoA-SH. The data for a high concentration of GlmU incubated in the presence of either 10 mM $MgCl_2$ or 1 mM EDTA, then used in the assay at 10 nM, using an assay buffer containing 10 mM $MgCl_2$ are represented by either ○ or ●, respectively.

The non-linear kinetics of GlmU acetyltransferase were further exacerbated by an initial incubation of the enzyme with 1 mM EDTA, a divalent metal chelator, in comparison to a similar incubation with 10 mM $MgCl_2$. These data confirmed the need for a new preparation of GlmU, purified and stored in the constant presence of 10 mM $MgCl_2$, was required to ensure optimal acetyltransferase activity, prior to progressing with detailed mechanistic studies.

3.3 Differential scanning fluorimetry assay configuration

CD thermal stability assays, although widely used for many years, have several limitations, including the requirement for high concentrations of protein and being a relatively low throughput method. To offer an alternative technique to monitor the integrity of GlmU secondary structure, in addition to addressing the limitations of using CD, a DSF assay was developed (see Section 135). This method can be miniaturised to utilise 384 well PCR microplates, making this approach amenable to studying ligand-induced protein stabilisation. Ligand-induced stabilisation is a phenomenon whereby the binding of ligands, such as inhibitors, substrates, metal ions, cofactors and other binding partners enhance the stability of a target protein due to the energetic coupling of the ligand-binding and target-melting reactions.

GlmU DSF using SYPRO Orange was carried out to monitor unfolding of the protein as the temperature of the sample is increased. To identify the optimal conditions for the DSF assay, GlmU was tested at four different concentrations in the presence of two different SYPRO Orange concentrations (Figure 31).

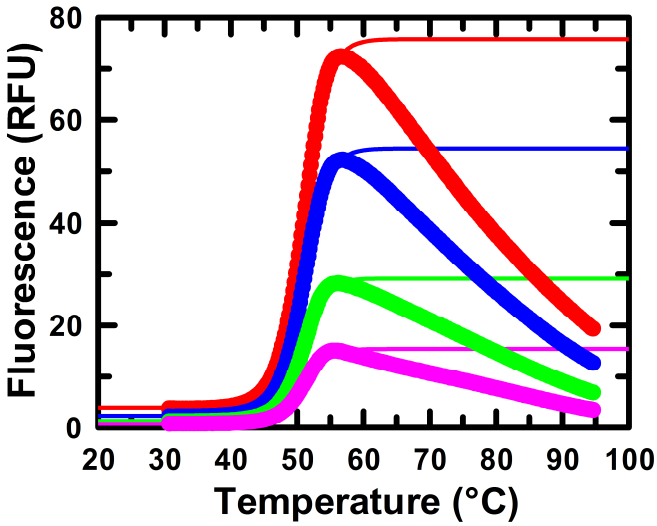


Figure 31. GlmU DSF assay optimisation.

Thermal melting of GlmU at the following concentrations; 8 µM (●), 4 µM (●), 2 µM (●) and 1 µM (●) in the presence of 5x SYPRO Orange, in assay buffer at pH 7.5, as the temperature is increased from 30 to 95 °C. These data are a mean of three discrete replicates that have been background subtracted using a buffer blank measurement. The data have been fitted, where appropriate, to Eqn. (19).

The data in Figure 31 is an example of the thermal stability profiles of four GlmU protein concentrations in the presence of 5x SYPRO Orange. Table 9 compares the GlmU T_m values for each different protein concentration, as well as the results of using two concentrations of SYPRO Orange.

Table 9. GlmU DSF assay optimisation.

[GlmU] (μM)	T_m (°C) in 10x SYPRO Orange [a, b]	Maximum fluorescence (RFU) in 10x SYPRO Orange	T_m (°C) in 5x SYPRO Orange [a, b]	Maximum fluorescence (RFU) in 5x SYPRO Orange
8	50.9 ± 0.03	57.4 ± 0.26	50.9 ± 0.02	75.8 ± 0.26
4	50.4 ± 0.04	30.8 ± 0.17	51.0 ± 0.02	54.4 ± 0.17
2	49.5 ± 0.05	11.9 ± 0.07	50.9 ± 0.03	29.2 ± 0.12
1	48.7 ± 0.06	5.1 ± 0.03	50.8 ± 0.04	15.4 ± 0.08

[a] T_m determined by monitoring SYPRO Orange fluorescence emission at 568 nm. [b] T_m determined in an assay buffer containing 25 mM HEPES, pH 7.5, 100 mM NaCl.

The data in Table 9 confirms, while optimising the DSF assay, that there is little difference in the apparent GlmU T_m regardless of the concentration of protein or SYPRO Orange used. The maximum fluorescence achieved for each combination of conditions explored is directly dependent on the concentration of GlmU used; a decrease in protein utilised leads to a reduction in the maximal fluorescence. The maximum fluorescence is increased by 1.3 to 3-fold as the concentration of SYPRO Orange is reduced from 10x to 5x. These experiments have identified the optimal combination of conditions that will be used for further DSF experiments.

3.4 Further protein preparation

Preliminary protein purification experiments and subsequent initial enzymatic characterisation experiments revealed a requirement of GlmU for Mg^{2+} to optimally catalyse acetyl transfer. Further protein preparation was carried out using His_6-TEV-GlmU construct in the continuous presence of 10 mM $MgCl_2$ to minimise any sub-optimal GlmU acetyltransferase activity. Apart from the continuous presence of 10 mM $MgCl_2$, there were other differences to the established purification methodology, including the addition of ion exchange (IEX) chromatography, followed by a second SEC step, to remove high concentrations of NaCl. The final product of the protein preparation process was analysed by SDS-PAGE gel electrophoresis (Figure 32), which highlights a single significant band, as well as by PMF and ESI-LC-MS.

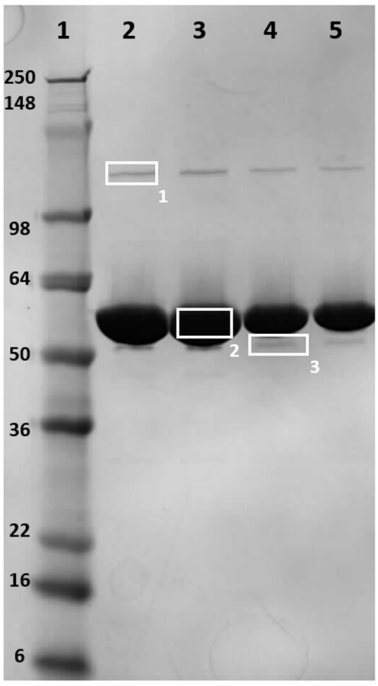


Figure 32. Gel electrophoresis of His_6-TEV-GlmU prepared in 10 mM $MgCl_2$.

SDS-PAGE analysis of the final purification product of His_6-TEV-GlmU, prepared in the presence of 10 mM $MgCl_2$. The gel layout is as follows; lane 1 is the gel MWt marker, lanes 2 and 3 are 10 µg load of protein and lanes 4 and 5 are 5 µg load of GlmU. Boxes 1 - 3 were the samples selected for PMF analysis.

The accurate protein mass was determined by ESI LC-MS as 53.4 kDa, which is consistent with N-terminal methionine deletion of His_6-TEV-GlmU. Protein concentration was measured as 2.2 mg/ml and a total of 66 mgs was purified from 30 grams of starting cell pellet. Protein identity was confirmed as GlmU using PMF analysis, which identified an average of 19 peptides after trypsin cleavage, from the highlighted bands in gel lanes 2, 3 and 4 (Figure 32, boxes 1-3), equating to 47 % coverage of the expected amino acid sequence. These data confirm the homogeneity, sequence and mass of the protein purified, in the continuous presence of 10 mM $MgCl_2$ from *E. coli* cell pellet as His6-TEV-GlmU.

The final pooled purification product of His_6-TEV-GlmU, prepared in the presence of 10 mM 10 mM $MgCl_2$ ($GlmU^{+Mg2+}$), was monitored by CD spectroscopy to confirm the presence of secondary polypeptide structure (Figure 33).

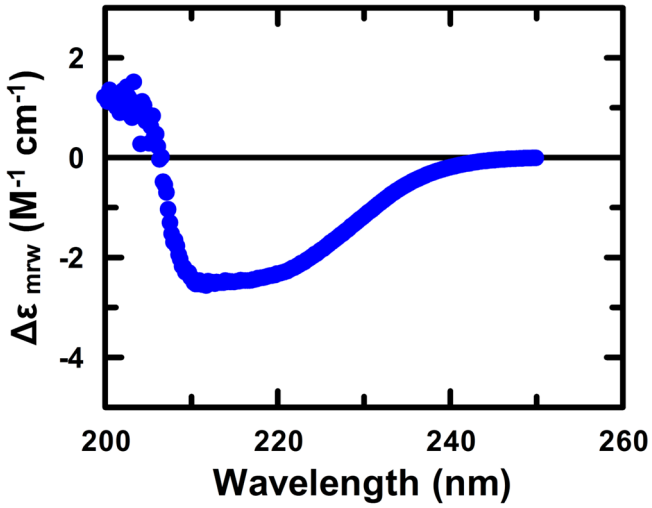


Figure 33. Far-UV CD spectra of GlmU^{+Mg2+}.

GlmU^{+Mg2+} protein secondary structure spectrum was determined by far-UV CD spectroscopy, using a wavelength scan from 190 – 250 nm. These data are a mean of three test occasions that have been background subtracted using a buffer blank measurement.

The far-UV CD spectra overlays with the CD spectra determined for His_6-TEV-GlmU purified in the absence of 10 mM $MgCl_2$ (GlmU^{-Mg2+}), which is typical of a significant proportion parallel β-helix (Figure 24)[418]. The GlmU^{+Mg2+} T_m was determined by using DSF, in the presence of 10 x SYPRO Orange, to monitor unfolding of protein as the temperature of the sample is increased. To provide a direct comparison, the same methodology was utilised to determine the T_m of GlmU^{-Mg2+} (Figure 34).

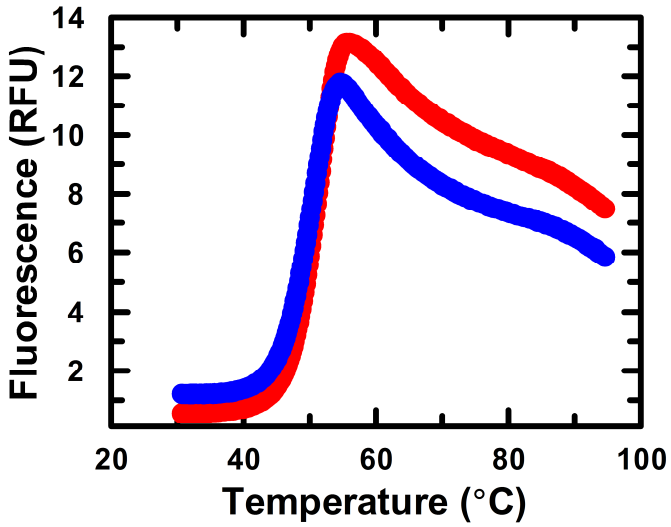


Figure 34. Comparison of the DSF melting profiles of GlmU^{+Mg2+} and GlmU^{-Mg2+}.

Thermal melting of GlmU^{+Mg2+} (●) and GlmU^{-Mg2+} (●) in the presence of 10 x SYPRO Orange, in assay buffer at pH 7.5, as the temperature is increased from 30 to 95 °C. These data are a mean of three discrete replicates that have been background subtracted using a buffer blank measurement. The data have been fitted, where appropriate, to Eqn. (19).

The T_m of GlmU^{+Mg2+} determined by DSF was 50.8 ± 0.2 °C, while the T_m of GlmU^{-Mg2+} was 49.5 ± 0.5 °C. The comparison of the T_m of the two preparations of GlmU demonstrates there is less than 1 °C difference, which indicates that the presence of Mg^{2+} has no significant effect on the stability of the secondary structure of the proteins. To confirm the DSF T_m results, alternative assessments of GlmU thermal melting were carried out using CD and NanoDSF (Table 10).

Table 10. Further T_m determination of GlmU^{+Mg2+} and GlmU^{-Mg2+}.

GlmU preparation	CD T_m (°C)[a, b]	NanoDSF T_m (°C)[b]
GlmU^{+Mg2+}	45.1 ± 0.3	49.3 ± 0.3
GlmU^{-Mg2+}	44.5 ± 0.1	47.8 ± 0.5

[a] T_m determined by monitoring CD of GlmU at 222 nm. [b] T_m determined using 2 µM GlmU in an assay buffer containing 25 mM HEPES, pH 7.5, 100 mM NaCl and 10 mM MgCl2 (depending on protein preparation used).

The data in Table 10 confirms the previous observations, using DSF, that there is a consistent, modest difference in T_m between GlmU^{+Mg2+} and GlmU^{-Mg2+}. The binding of Mg^{2+} is believed to have a structural role and these data indicate, that in the presence of 10 mM of the cation, the structure is modestly stabilised. A comparison of the apparent GlmU thermal stability data determined by the three different methodologies highlights that the CD T_m values are between 4 - 5 °C lower than the protein melting values determined by DSF and NanoDSF. DTP absorbance assay was utilised to evaluate the acetyltransferase activity of both GlmU^{-Mg2+} and GlmU^{+Mg2+} by kinetically monitoring the formation of CoA-SH in an assay buffer containing 10 mM $MgCl_2$ at pH 7.5, without any pre-incubations with 10 mM $MgCl_2$ (Figure 35).

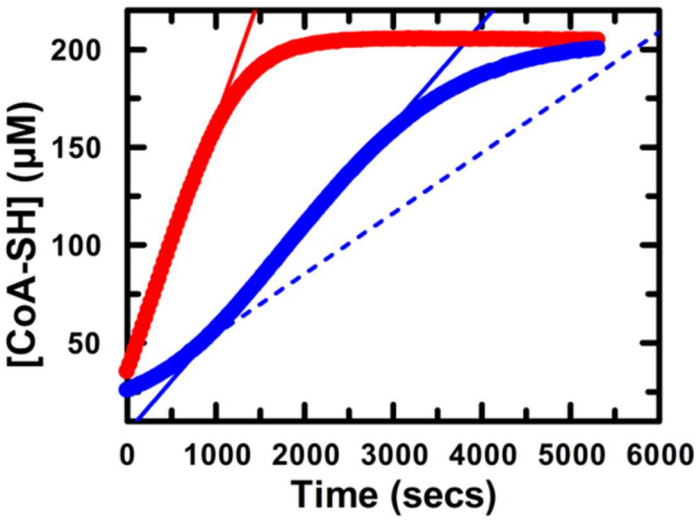


Figure 35. Acetyltransferase activity of GlmU^{+Mg2+} and GlmU^{-Mg2+} preparations.
Progress curves of 10 nM GlmU^{+Mg2+} (●) and GlmU^{-Mg2+} (●) acetyltransferase reactions using DTP to monitor formation of CoA-SH at pH 7.5 every 10 secs. The progress curves for GlmU^{+Mg2+} were fit to a linear regression model for time points between 0 to 800 secs (—). The progress curves for GlmU^{-Mg2+} were fit to two linear regression models (1) for time points between 0 to 650 secs (—) and (2) for time points between 1000 to 3000 secs (---). These data are a mean of four discrete replicates that have been background subtracted using a buffer blank measurement.

The non-linear kinetics initially observed for GlmU acetyltransferase activity in the absence of added $MgCl_2$ are eradicated when using the GlmU^{+Mg2+} preparation of enzyme (Figure 35). As a control, and comparator, it is evident the non-linear acetyltransferase activity is still present when using the GlmU^{-Mg2+} preparation of enzyme. appear to be eradicated by. In contrast, the progress plot for 10 nM GlmU^{-Mg2+} has two phases that were fit to two linear rates; the first was 0.03 CoA-SH s^{-1} and the second was 0.05 CoA-SH s^{-1}, whereas the rate of GlmU^{+Mg2+} acetyl transfer was 0.12 CoA-SH s^{-1}. These data confirmed the requirement for 10 mM $MgCl_2$, from protein purification to assay, to eradicate the non-linear GlmU acetyltransferase kinetics previously observed using the thiol detection assays. However, the presence of $MgCl_2$ only modestly effected the CD profile and the thermal stability of GlmU. To better understand the need for $MgCl_2$ for optimal GlmU acetyltransferase kinetics and whether this requirement is concentration dependent, divalent metal ion was titrated in the presence of fixed concentrations of Ac-CoA and GlcN-1P and CoA-SH release monitored using the DTP absorbance assay (Figure 36).

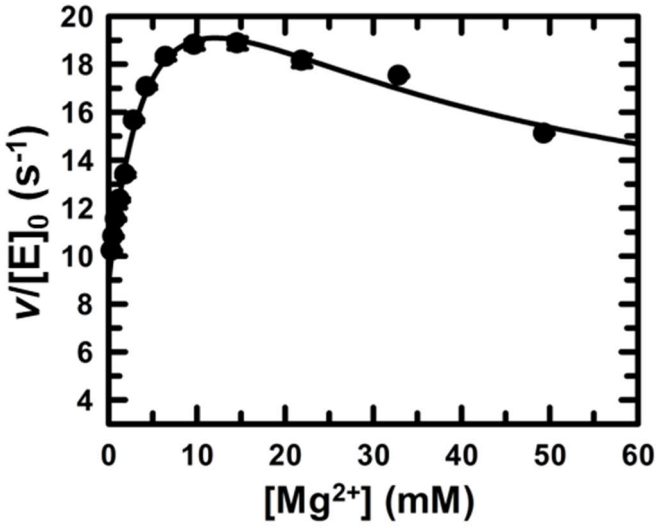


Figure 36. GlmU^{-Mg2+} acetyltransferase apparent Mg^{2+} K_{act}

Titration of $MgCl_2$ in the presence of 5 nM GlmU^{-Mg2+}, 400 µM DTP and 200 µM Ac-CoA and GlcN-1P. The initial velocity for each condition was determined and subsequently correlated with the Mg^{2+} concentration range utilised and fitted to Eqn. (2). Symbols represent experimental data, and solid lines the fit of the data to the appropriate model. These data are a mean of 3 discrete replicates that have been background subtracted using a minus GlmU control measurement.

Monitoring the acetyltransferase activity of GlmU^{-Mg2+} in the presence of a titration of $MgCl_2$ enabled determination of an apparent $K_{act,\ Mg2+}$ of 5.7 ± 1.3 mM and identification of a Mg^{2+} concentration for optimal acetyl transfer of 10 mM (Figure 36). This experiment was carried out at 200 µM concentrations for both Ac-CoA and GlcN-1P, which are equivalent to the $K_{m,\ app}$ values reported in the literature. In addition, the use of the GlmU^{-Mg2+} preparation could be flawed as there was a possibility that some divalent metal ion could be bound to the enzyme. To better understand the GlmU acetyltransferase activity, it was first necessary to determine the steady state kinetic parameters for Ac-CoA and GlcN-1P in the presence of 10 mM Mg^{2+}. Furthermore, to fully understand whether the requirement for Mg^{2+} is either specific or concentration dependent, a variety of divalent metal ions will be investigated over a range of concentrations. Preparations of GlmU treated with EDTA to remove any residual divalent metal ion, will be utilised to carry out both specificity and concentration studies.

3.5 Uridylyltransferase assay configuration

The uridylyltransferase activity of GlmU was monitored using Pi assays to detect the accumulation of the product PPI in real time. The two detection methodologies employed to detect Pi, which is formed by PPase from the uridylyltransferase product PPi, were both fluorogenic reagents; either the fluorescent Pi binding protein MDCC-PBP or a multi-enzyme coupled system, referred to as PiXan, that ultimately generates the fluorescent product, resorufin. The two methodologies are sensitive to levels of free Pi in solution and are dependent on the conversion of PPi by PPase, which is utilised at high concentrations. The rationale for the utilisation of high levels of PPase is to prevent the enzymatic conversion of PPi to Pi from becoming the rate-limiting catalytic step.

To initially compare the sensitivity of MDCC-PBP and PiXan, Pi was titrated in assay buffer at pH 7.5, prior to addition of the detection reagent (Figure 37).

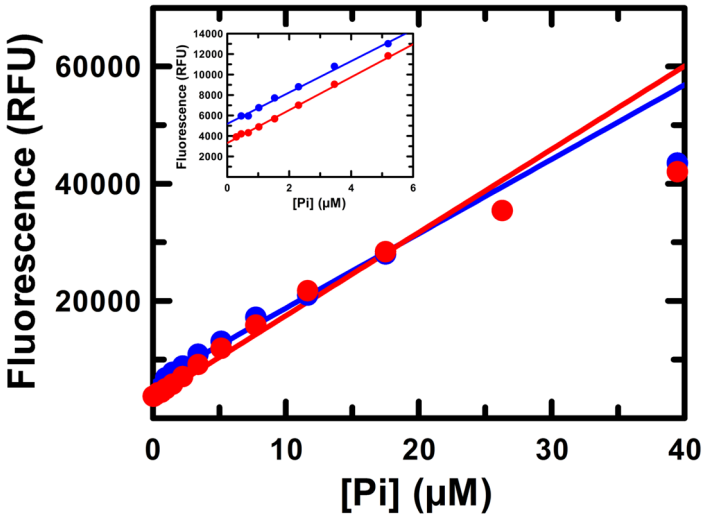


Figure 37. Detection of Pi standard titrations.

Pi standard was titrated in assay buffer at pH 7.5 and subsequently detected by additions of 5 μM of either MDCC-PBP (●) or PiXan reagents (●). Fluorescence, measured as relative fluorescence units (RFU) was monitored at emission wavelengths of either 470 nm (MDCC-PBP) or 580 nm (PiXan). Each of these data have been fitted to a linear regression; MDCC-PBP (—, correlation coefficient = 0.99) or PiXan reagents (—, correlation coefficient = 0.99). Inset figure, Pi concentration range from 0.5 to 6 μM.

The use of either MDCC-PBP and PiXan reagents to detect the Pi were equivalent at pH 7.5, with a minimum level of detection of 0.5 μM that was linear up to a concentration of 20 μM Pi.

To further investigate and evaluate the Pi detection methodologies, MDCC-PBP and PiXan reagents were used to monitor the uridylyltransferase activity of GlmU by kinetically detecting the formation of PPi in assay buffer at pH 7.5 (Figure 38).

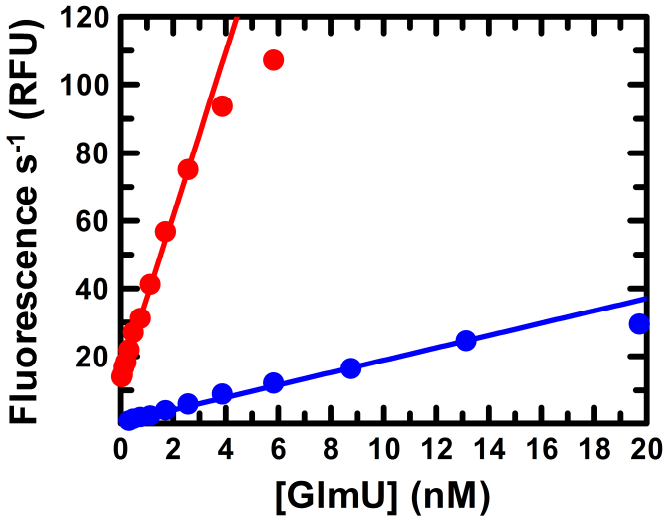


Figure 38. Validation of Pi detection assays to monitor GlmU uridylyltransferase activity.

GlmU titration in the presence of fixed concentrations of 100 µM GlcNAc-1P and UTP in assay buffer at pH 7.5 and subsequently detected by additions of 5 µM of either MDCC-PBP (●) or PiXan reagents (●). Fluorescence, measured as RFU s^{-1} was monitored at emission wavelengths of either 470 nm (MDCC-PBP) or 580 nm (PiXan). Each of these data have been fitted to a linear regression; MDCC-PBP (—, correlation coefficient = 0.99) or PiXan reagents (—, correlation coefficient = 0.99).

There was difference between the sensitivity and the total fluorescence of the Pi detection assays to monitor the release of PPi as it is formed by GlmU uridylyltransferase activity. The MDCC-PBP assay generates less fluorescence per second but the methodology was able to monitor a linear increase of PPi, per concentration of GlmU utilised up to 15 nM enzyme. The PiXan assay generates more fluorescence per second but the methodology appears to be less sensitive, in terms of linearity of PPi produced per concentration of GlmU utilised up to 4 nM. The apparent lack of sensitivity of the PiXan methodology is potentially due to the generation of high fluorescence signal, which has been observed to saturation the microplate reader photo-multiplier tube (PMT) detector. The saturation of the PMT detector can occur rapidly, especially at higher concentrations of GlmU, which leads to difficulties in determining an accurate initial rate.

These data confirmed that both uridylyltransferase, Pi detection methodologies are fit for purpose and can be used to monitor the GlmU catalysed generation of the product PPi.

3.6 Acetyltransferase steady state kinetics

To initially characterise the acetyltransferase activity of GlmU (utilising $GlmU^{+Mg2+}$) titrations of either GlcN-1P or Ac-CoA in the presence of saturating concentrations of the second substrate, were investigated. These experiments were carried out in an assay buffer containing 10 mM $MgCl_2$ and CoA-SH product formation was detected kinetically using the DTP absorbance assay. Initial velocities were calculated for each test condition and the data were subsequently fitted to determine the acetyltransferase Michaelis-Menten kinetic parameters (Figure 39).

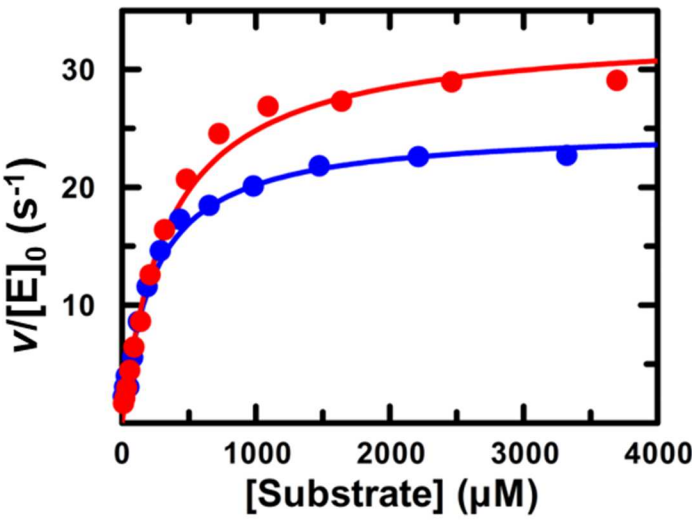


Figure 39. Acetyltransferase steady state kinetics.

Initial velocity data for GlmU acetyltransferase reaction using titrations of either Ac-CoA (●) or GlcN-1P (●). Symbols represent experimental data, and solid lines the fit of the data to Eqn. (1), for either Ac-CoA (—) or GlcN-1P (—). These data are a mean of four discrete replicates.

The acetyltransferase activity of $GlmU^{+Mg2+}$ was monitored in the presence of a titration of Ac-CoA, at a saturating concentration of GlcN-1P, which enabled determination of a $K_{m,\ Ac\text{-}CoA}$ of 240.1 ± 20 μM and $k_{cat,\ Ac\text{-}CoA}$ = 25.1 ± 0.6 s^{-1}. The reverse experiment, using a titration of GlcN-1P in the presence of a high concentration of Ac-CoA was also fitted to Eqn. (1), giving a $K_{m,\ GlcN\text{-}1P}$ of 344.8 ± 29.6 μM and $k_{cat,\ GlcN\text{-}1P}$ = 33.3 ± 0.9 s^{-1}. The k_{cat} is independent of substrate varied

and the variation between the values determined by fitting the GlmU acetyltransferase substrate titrations is likely due to non-saturating conditions utilised for Ac-CoA. The GlmU acetyltransferase Michaelis-Menten parameters will be used to further investigate and quantify the requirement for divalent metal ions for optimal activity.

3.7 Uridylyltransferase steady state kinetics

To initially characterise the uridylyltransferase activity of GlmU (utilising $GlmU^{+Mg2+}$), titrations of either GlcNAc-1P or UTP in the presence of saturating concentrations of the second substrate, were investigated. These experiments were carried out in an assay buffer containing 10 mM $MgCl_2$ and the use of either MDCC-PBP or PiXan assay methodologies to monitor PPi product formation. Initial velocities were determined for each test condition and the data were subsequently fitted to determine the uridylyltransferase Michaelis-Menten kinetic parameters (Figure 40).

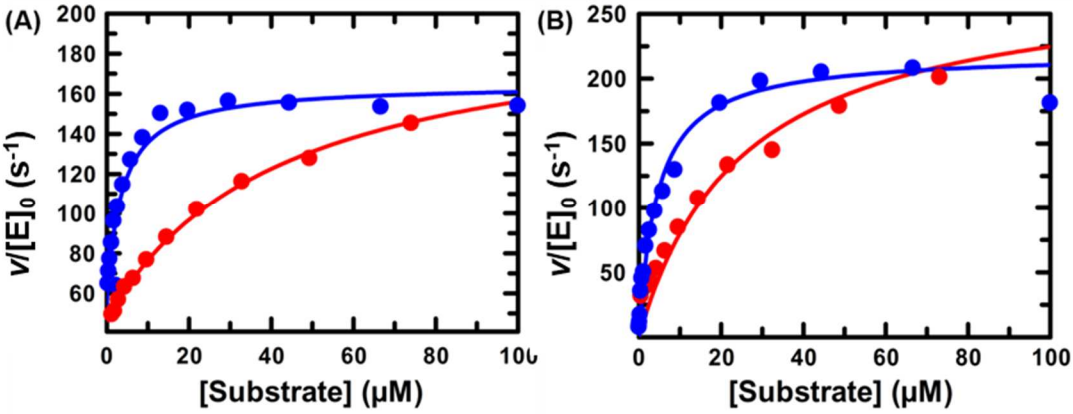


Figure 40. Uridylyltransferase steady state kinetics.

(A) Initial velocity data for GlmU uridylyltransferase reaction monitored using a PiXan assay. (B) Initial velocity data for GlmU uridylyltransferase reaction monitored using a MDCC-PBP assay. Data generated using titrations of either UTP (●) or GlcNAc-1P (●). Symbols represent experimental data, and solid lines the fit of the data to Eqn. (1), for either UTP (—) or GlcNAc-1P (—). These data are a mean of five discrete replicates.

The uridylyltransferase activity of $GlmU^{+Mg2+}$ was monitored using PiXan in the presence of a titration of UTP, at a saturating concentration of GlcNAc-1P, which enabled determination of a $K_{m,\ UTP}$ of 3.6 ± 1.3 μM and $k_{cat,\ UTP}$ = 150.1 ± 9.2 s^{-1}. The same experiment was carried out using MDCC-PBP, which yielded a $K_{m,\ UTP}$ of 4.5 ± 0.5 μM and $k_{cat,\ UTP}$ = 219.9 ± 7.4 s^{-1}.

The reverse experiment, titrating GlcNAc-1P in the presence of a high concentration of UTP and monitored using PiXan was also fitted to Eqn. (1), giving a $K_{m,\ GlcNAc\text{-}1P}$ of 41.8 ± 2.8 μM and $k_{cat,\ GlcNAc\text{-}1P}$ = 155.4 ± 2.7 s^{-1}. The same experiment was carried out using MDCC-PBP, which yielded a $K_{m,\ GlcNAc\text{-}1P}$ of 25.1 ± 3.2 μM and $k_{cat,\ GlcNAc\text{-}1P}$ = 221.2 ± 8.4 s^{-1}.

The variation between the k_{cat} values determined by fitting the GlmU uridylyltransferase substrate titrations data is likely due to non-saturating conditions utilised for the second substrate. However, these k_{cat} variations maybe due to the Pi detection assays used, primarily the difference in sensitivity, in addition to the relative complexity of PiXan reagents.

3.8 Divalent metal ion modulation of acetyl transfer

To better understand the Mg^{2+} requirement for optimal GlmU acetyltransferase activity, it was necessary to treat both $GlmU^{+Mg2+}$ and $GlmU^{-Mg2+}$ preparations with EDTA, prior to characterising the effect of Mg^{2+}. This EDTA treatment was carried out to remove any residual, tightly bound, divalent metal ions that might not have diffused from the protein during the purification process.

The two preparations of GlmU were incubated with 10 mM EDTA, in a buffer comprising of 25 mM HEPES, pH 7.5, 100 mM NaCl, at 4 °C for 18 hours. This incubation was followed by dialysis for 24 hours (buffer change after 12 hours) at 4 °C while constantly stirring to remove the 10 mM EDTA. The EDTA-treated and dialysed protein preparations, known hereon in as $GlmU^{+Mg2+,\ EDTA}$ and $GlmU^{-Mg2+,\ EDTA}$, were investigated to characterise both the thermal melting profiles and the acetyltransferase activity (Figure 41).

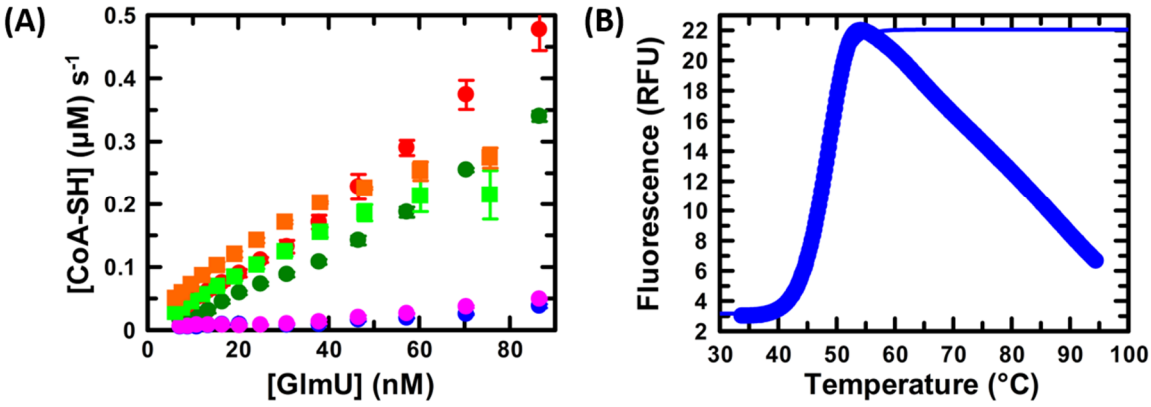


Figure 41. DSF stability and acetyltransferase activity evaluation of EDTA-treated GlmU.

(A) Titrations of GlmU to monitor acetyltransferase activity and determine the initial velocity in the presence of 200 μM substrates, as [CoA-SH] s^{-1}, of the following preparations; (●) $GlmU^{-Mg2+,\ EDTA}$, (●) $GlmU^{-Mg2+,\ EDTA}$ in the presence of 10 mM Mg^{2+}, (●) $GlmU^{+Mg2+,\ EDTA}$, (●) $GlmU^{+Mg2+,\ EDTA}$ in the presence of 10 mM Mg^{2+}, (■) $GlmU^{-Mg2+,\ EDTA}$ in the presence of 10 mM Mn^{2+} and (■) $GlmU^{+Mg2+,\ EDTA}$ in the presence of 10 mM Mn^{2+}. (B) Thermal melting of EDTA-treated $GlmU^{-Mg2+}$ (●) in the presence of 10 x SYPRO Orange, in assay buffer at pH 7.5, as the temperature is increased from 30 to 95 °C. These data are a mean of 16 discrete replicates that have been background subtracted using a buffer blank measurement. The thermal profile data have been fitted, where appropriate, to Eqn. (19).

The data in Figure 41 demonstrates $GlmU^{-Mg2+,\ EDTA}$ had reduced acetyltransferase activity, when compared to the WT enzyme. The acetyltransferase activity of the $GlmU^{-Mg2+,\ EDTA}$ in an assay buffer including either 10 mM $MgCl_2$ or $MnCl_2$ was increased by 5 – 10-fold, from 0.005, in the absence of cation, to a maximum of 0.05 CoA-SH s^{-1} for 10 nM enzyme, with equivalent activities to WT GlmU in the same assay buffer ($GlmU^{-Mg2+}$ in buffer containing 10 mM $MgCl_2$). The acetyltransferase activity of $GlmU^{+Mg2+,\ EDTA}$ in assay buffer including either 10 mM $MgCl_2$ or $MnCl_2$ was increased by 5 – 10-fold, to a level that is equivalent to WT GlmU. The GlmU acetyltransferase activity for all protein preparations, in the absence of 10 mM Mg^{2+} had an initial, non-linear phase. The DSF thermal stability profile of $GlmU^{-Mg2+,\ EDTA}$ was fitted to determine an apparent T_m = 47.9 ± 0.1 °C, which was 2 °C lower than the T_m for $GlmU^{-Mg2+}$ (Table 9).

DSF was then used to characterise the effect of titrating $MgCl_2$ from 10 mM on the thermal melting of $GlmU^{-Mg2+,\ EDTA}$ Figure 42.

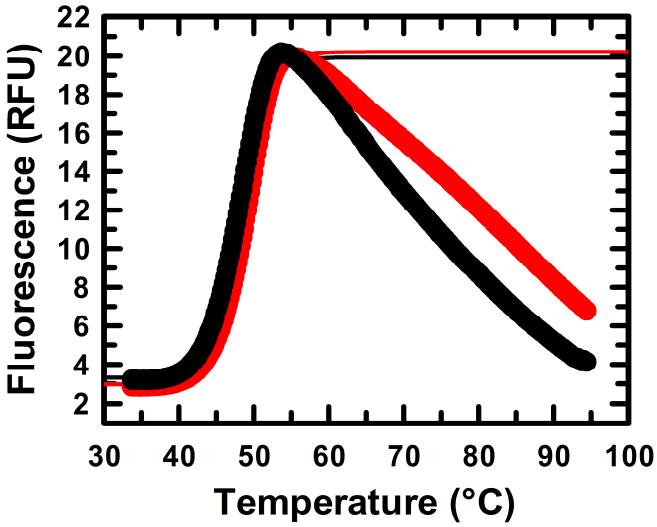


Figure 42. The effect of Mg^{2+} on $GlmU^{-Mg2+,\ EDTA}$ T_m.

(A) DSF thermal stability profile of $GlmU^{-Mg2+,\ EDTA}$ in the presence of either 10 mM (●) or 0 mM (●) Mg^{2+}. These data are a mean of 3 discrete replicates that have been background subtracted using a minus GlmU control measurement. The data have been fitted, where appropriate, to a Eqn. (19).

The DSF T_m of $GlmU^{-Mg2+,\ EDTA}$ was stabilised by 1.2 °C in the presence of 10 mM Mg^{2+} in Figure 42. The $GlmU^{-Mg2+,\ EDTA}$ DSF T_m was stabilised in a concentration dependent manner from 47.71 ± 0.05 °C to 48.91 ± 0.04 °C. To characterise the apparent activation of GlmU acetyltransferase activity by Mg^{2+}, the divalent metal ion was titrated in the presence of 200 µM Ac-CoA and GlcN-1P, and CoA-SH release was monitored using the DTP absorbance assay (Figure 43).

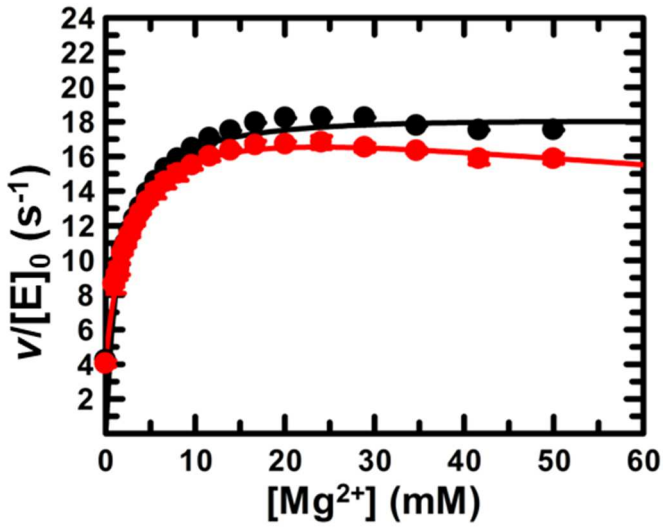


Figure 43. GlmU acetyltransferase apparent Mg^{2+} $K_{act.}$

Titration of Mg^{2+} in the presence of either 12 nM $GlmU^{-Mg2+,\ EDTA}$ (●) or $GlmU^{+Mg2+,\ EDTA}$ (●), 400 µM DTP and 200 µM Ac-CoA and GlcN-1P. The initial velocity for each

condition was determined and subsequently correlated with the Mg^{2+} concentration range utilised and fitted to Eqn.(2). Symbols represent experimental data, and solid lines the fit of the data to the appropriate model. These data are a mean of 6 discrete replicates.

Monitoring the acetyltransferase activity of two different preparations of enzyme, either $GlmU^{-Mg2+, EDTA}$ or $GlmU^{+Mg2+, EDTA}$ in the presence of a titration of Mg^{2+}, yielded apparent $K_{act, Mg2+}$ values of either 1.40 ± 0.15 mM or 1.16 ± 0.13 mM, respectively. The concentration of Mg^{2+} required for optimal acetyl transfer was determined as 10 mM (Figure 43). Activation of GlmU V_{max} by Mg^{2+} is in the range of 4 to 4.5-fold, however this experiment was carried out in non-saturating concentrations of both Ac-CoA and GlcN-1P. Further studies were carried out to understand the dependence of GlmU acetyltransferase Michaelis-Menten parameters on Mg^{2+} concentration. Titrations of either Ac-CoA or GlcN-1P were prepared in the presence of high concentrations of the second substrate and varying concentrations of Mg^{2+}. $GlmU^{-Mg2+, EDTA}$ was added to the saturation titrations and acetyltransferase activity was monitored using the DTP absorbance assay (Figure 44).

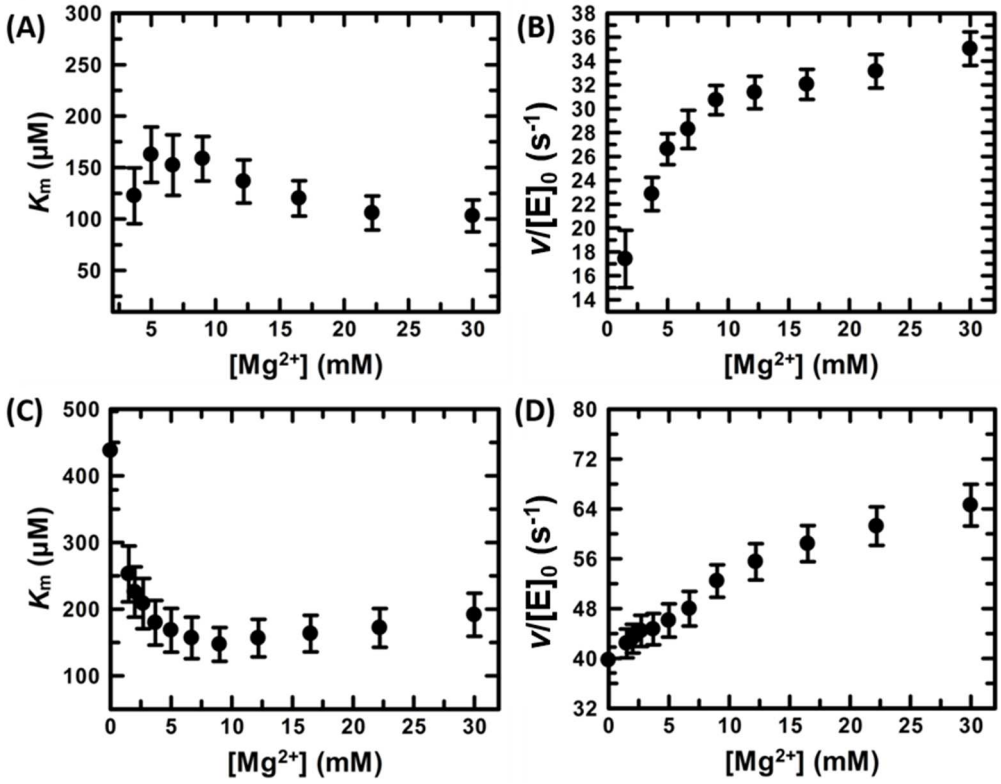


Figure 44. Effect of $MgCl_2$ on GlmU acetyltransferase steady-state kinetics.

$GlmU^{-Mg2+, EDTA}$ acetyltransferase activity monitored detection of CoA-SH using the DTP absorbance assay. (A-B) Titration of Ac-CoA at saturating concentrations of GlcN-1P and varying concentrations of Mg^{2+}. (A) Effect of varying Mg^{2+}

concentrations on $K_{m,\ Ac\text{-}CoA}$, (B) Effect of varying Mg^{2+} concentrations on $k_{cat,\ Ac\text{-}CoA}$. (C-D) Titration of GlcN-1P at saturating concentrations of Ac-CoA and varying concentrations of Mg^{2+}. (C) Effect of varying Mg^{2+} concentrations on $K_{m,\ GlcN\text{-}1P,}$ (D) Effect of varying Mg^{2+} concentrations on $k_{cat.\ GlcN\text{-}1P}$. Symbols represent a mean of three replicates, which were subsequently fitted to either Eqn. (1) or (3).

The dependence of GlmU acetyltransferase kinetic parameters on Mg^{2+} was only modest, with little effect observed on either $K_{m,\ Ac\text{-}CoA}$ or $K_{m,\ GlcN\text{-}1P}$ (Figure 44 A and C). A greater effect was observed on both k_{cat} values with increases of 1.5 to 2-fold at high concentrations of metal ion, confirming the observations in Figure 43 (Figure 44 B and D). The data in Figure 44 can be re-plotted to define the effect of varying the acetyltransferase substrates on the $K_{act,\ Mg2+}$ (Figure 45).

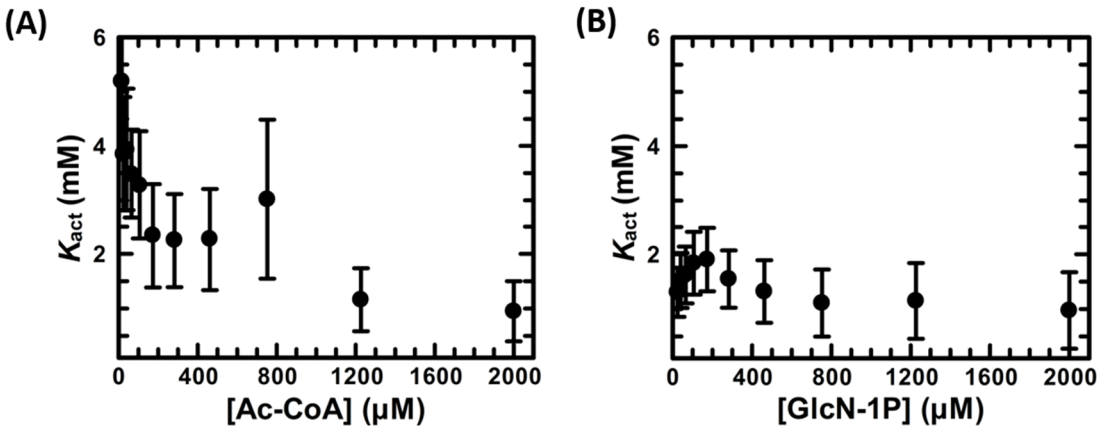


Figure 45. Effect of varying acetyltransferase substrates on $K_{act,\ Mg2+}$.

Re-plot of the data in Figure 44. (A) Titration of Mg^{2+} at saturating concentrations of GlcN-1P and varying concentrations of Ac-CoA. Effect of varying Ac-CoA concentrations on $K_{act,\ Mg2+.}$ (B) Titration of Mg^{2+} at saturating concentrations of Ac-CoA and varying concentrations of GlcN-1P. Effect of varying GlcN-1P concentrations on $K_{act,\ Mg2+}$. Symbols represent experimental data fitted to either Eqn. (1) or (3).

The plots in Figure 45 demonstrate that GlmU acetyltransferase activity $K_{act,\ Mg2+}$ varies by approximately 2 mM as the concentration of either Ac-CoA or GlcN-1P are changed. These data define the requirement of GlmU for Mg^{2+} to achieve optimal acetyltransferase activity.

To further understand the divalent metal ion requirement of GlmU for optimal acetyltransferase activity, several cations were investigated, including Mn^{2+}, Ca^{2+}, Ni^{2+}, Co^{2+} and Zn^{2+}. Initially, the divalent metal ion effect on GlmU DSF T_m was investigated by titrating the test cations in the presence of fixed concentration of

GlmU$^{-Mg2+,\ EDTA}$ and SYPRO Orange, prior to application of a temperature gradient (Table 11).

Table 11. Divalent metal ion effect on GlmU T_m.

Cation	GlmU DSF T_m (°C) 10 mM cation [a, b]	GlmU DSF T_m (°C) 0.6 mM cation [a, b, c]	Maximum fluorescence (RFU) 10 mM cation [c]	Maximum fluorescence (RFU) 0.6 mM cation [c, d]
Mg^{2+}	48.87 ± 0.04	47.71 ± 0.05	20.1 ± 0.2	22.1 ± 0.1
Mn^{2+}	51.71 ± 0.15	50.30 ± 0.04	16.5 ± 0.3	17.5 ± 0.4
Ca^{2+}	47.95 ± 0.04	47.87 ± 0.05	14.8 ± 0.1	22.2 ± 0.2
Co^{2+}	49.91 ± 0.08	47.87 ± 0.04	3.5 ± 0.2	4.5 ± 0.4
Ni^{2+}	52.16 ± 0.11	48.69 ± 0.03	9.1 ± 0.1	12.8 ± 0.2
Zn^{2+} [e]	N.D.	N.D.	26.7 ± 0.2	19.2 ± 0.5

[a] T_m determined by monitoring SYPRO Orange fluorescence emission at 568 nm. [b] T_m determined in an assay buffer containing 25 mM HEPES, pH 7.5, 100 mM NaCl. Thermal melting curves were fitted to Eqn. (19). [c] The minimum concentration of test cation was 0.6 mM. [d] Peak fluorescence (RFU) of the thermal melting profile. [e] Elevated fluorescence of between approximately 16 and 28 RFU at 35 °C.

The DSF T_m of GlmU$^{-Mg2+,\ EDTA}$ in the presence of increasing concentrations of Mn^{2+} was stabilised up to a maximum of approximately 4 °C at a concentration of 10 mM. The maximum fluorescence of the thermal melting curve in the presence of 10 mM Mn^{2+} was 16.5 RFU, which is a reduction of approximately 30% in comparison to the thermal profile in the absence of cation (Table 11). The DSF T_m of GlmU$^{-Mg2+,\ EDTA}$ was not stabilised at any test concentration of Ca^{2+}. The maximum fluorescence of the thermal melting curve in the presence of 10 mM Mn^{2+} was 14.8 RFU, which is a reduction of approximately 33% in comparison to the minus cation thermal profile (Table 11).

The DSF T_m of GlmU$^{-Mg2+,\ EDTA}$ appeared to be stabilised by both 10 mM Co^{2+} and Ni^{2+} by 2.2 °C or 4.4 °C, respectively. However, the maximum fluorescence of the thermal melting curve in the presence of either 10 mM Co^{2+} or Ni^{2+} were reduced by approximately 85% or 59% in comparison to the minus cation thermal profile. The

total SYPRO Orange fluorescence is dependent on the concentration of protein present, as the fluorophore is only fluorescent when bound to the hydrophobic patches revealed upon temperature induced unfolding. At higher temperatures, when protein is unfolded non-specific aggregation occurs occluding the hydrophobic patches and a subsequent decrease in fluorescence is observed (see Figure 42 for a typical protein thermal profile). The reduction in total fluorescence in the presence of 10 mM Co^{2+} and Ni^{2+} is likely due to protein aggregation, prior to application of the temperature gradient, which would reduce the total concentration of protein that can be monitored in the DSF experiment. The Zn^{2+} DSF experiments were not valid, as in contrast to the Co^{2+} and Ni^{2+} experiments, the total fluorescence was increased. The Zn^{2+} thermal profiles were distinct, as the starting fluorescence, at approximately 30 °C, were greater than 16 RFU, which is likely to be indicative of partial unfolding of GlmU, in contrast to the aggregation caused by Co^{2+} and Ni^{2}. The aggregation and unfolding caused by Co^{2+}, Ni^{2+} and Zn^{2+} will make subsequent activation experiments difficult to interpret, so no further studies were attempted with these cations.

To investigate whether either Mn^{2+} or Ca^{2+} alleviate the lag phase of GlmU-catalysed acetyl transfer, experiments were carried out to monitor initial acetyltransferase activity of $GlmU^{-Mg2+,\ EDTA}$ in the presence of 10 mM of either cation (Figure 46).

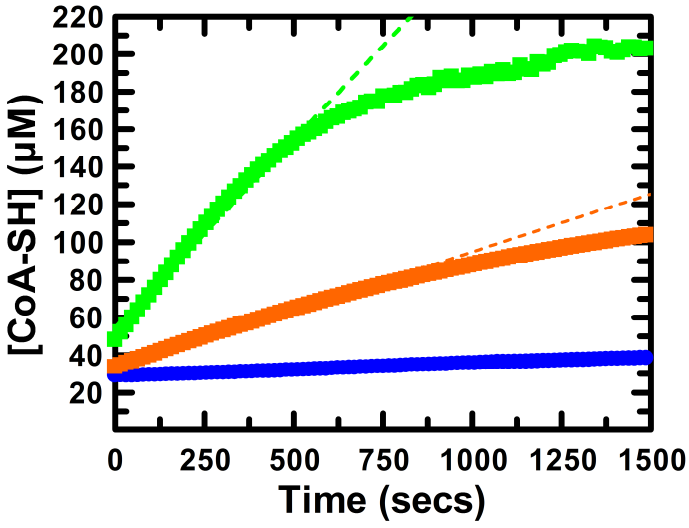


Figure 46. GlmU acetyltransferase activity evaluation with either 10 mM Mn^{2+} or Ca^{2+}.

GlmU acetyltransferase activity progress curves in the presence of either 10 mM Mn^{2+} (■) or Ca^{2+} (■). The acetyltransferase activity progress curves were monitored using 10 nM $GlmU^{-Mg2+,\ EDTA}$, 200 μM Ac-CoA and GlcN-1P using 400 μM DTP to

monitor formation of CoA-SH every 16 secs. The progress curves in the presence of 10 mM Mn^{2+} or Ca^{2+} were fit to a linear regression model for time points between 0 to 750 secs (---) and 0 to 400 secs (---), respectively. Control progress curves were monitored for 10 nM $GlmU^{-Mg2+, EDTA}$ (●) in the absence of cation. These data are a mean of six discrete replicates that have been background subtracted using a buffer blank measurement.

The data in Figure 46 demonstrated that the addition of 10 mM Mn^{2+} led to an approximate ten-fold increase in GlmU acetyltransferase activity, from 0.006 to 0.06 [CoA-SH] s^{-1}. The addition of 10 mM Ca^{2+} led to an approximate twenty-five-fold increase in GlmU acetyltransferase activity, from 0.006 to 0.18 [CoA-SH] s^{-1}. The GlmU acetyltransferase activity lag phase was alleviated in the presence of both 10 mM Mn^{2+} and Ca^{2+}.

The apparent Mn^{2+} and Ca^{2+} activation of GlmU-catalysed acetyl transfer was characterised by titrating the cations in the presence of $GlmU^{-Mg2+, EDTA}$ and 200 µM Ac-CoA and GlcN-1P. The acetyltransferase activity was monitored using the DTP absorbance assay to detect CoA-SH formation (Figure 47).

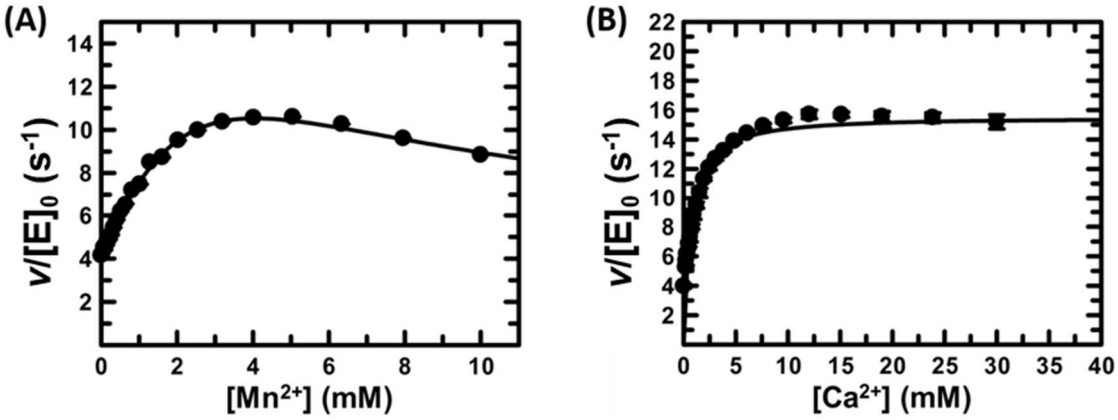


Figure 47. GlmU acetyltransferase apparent Mn^{2+} and Ca^{2+} K_{act}.

Titration of either Mn^{2+} (A) or Ca^{2+} (B) in the presence of 12 nM $GlmU^{-Mg2+, EDTA}$, 400 µM DTP and 200 µM Ac-CoA and GlcN-1P. The initial velocity for each condition was determined and subsequently correlated with the cation concentration range utilised and fitted to Eqn. (2). Symbols represent experimental data, and solid lines the fit of the data to the appropriate model. These data are a mean of 3 discrete replicates.

Table 12. GlmU acetyltransferase steady state parameters for activation by divalent metal ions.

Cation	K_{act} (mM) [a, b]	V_{max} (s^{-1}) [a, b]	V/K ($M^{-1}s^{-1}$)
Mg^{2+}	1.40 ± 0.15	18.70 ± 0.42	1.3 ± 0.28 x 10^4
Mn^{2+}	2.62 ± 0.62	10.30 ± 0.35	3.9 ± 0.56 x 10^3
Ca^{2+}	0.65 ± 0.08	15.71 ± 0.56	2.4 ± 0.7 x 10^4

[a] K_{act} determined by titrating divalent metal ions in the presence of 12 nM $GlmU^{-Mg2+, EDTA}$ and 200 µM Ac-CoA and GlcN-1P. CoA-SH formation monitored by DTP absorbance change at 324 nm. [b] Data fitted to either Eqn. (1) or (2), which was utilised for the Mn^{2+} titration.

The data in Figure 47 and Table 12 are the steady state parameters for divalent metal ion activation of GlmU-catalysed acetyl transfer. The activation of GlmU acetyltransferase activity V_{max} by Mn^{2+} was approximately 2.5-fold, were as V_{max} activation by Ca^{2+} was 4-fold. The concentration of either Mn^{2+} or Ca^{2+} required for optimal GlmU acetyltransferase activity was approximately either 4 mM or 6 mM, respectively.

These data demonstrate that divalent metal ion is essential for GlmU acetyltransferase activity and, after treatment with EDTA, little activity is observed. Further studies have highlighted that a range of divalent metal ions have both an additive and antagonistic effect on acetyltransferase activity. These findings will be discussed further in later sections but, for the further studies presented in Chapters 3 and 4, all experiments will be carried out in assay buffers containing 10 mM $MgCl_2$ and the will utilise the $GlmU^{+Mg2+}$ preparation (referred to from hereon in as “GlmU”) unless otherwise stated. The selection of 10 mM $MgCl_2$ in preference to the other divalent metals studied in this and the previous section are twofold; (1) Mg^{2+} is likely to be essential for uridylyltransferase activity measurements, and (2) large scale protein production had already been prepared in the presence of 10 mM $MgCl_2$.

3.9 Acetyltransferase initial velocity patterns

To characterise the acetyltransferase kinetic mechanism, initial velocity studies were carried out by varying Ac-CoA concentrations at various fixed GlcN-1P concentrations, in an assay buffer at pH 7.5, containing 10 mM $MgCl_2$ (Figure 48).

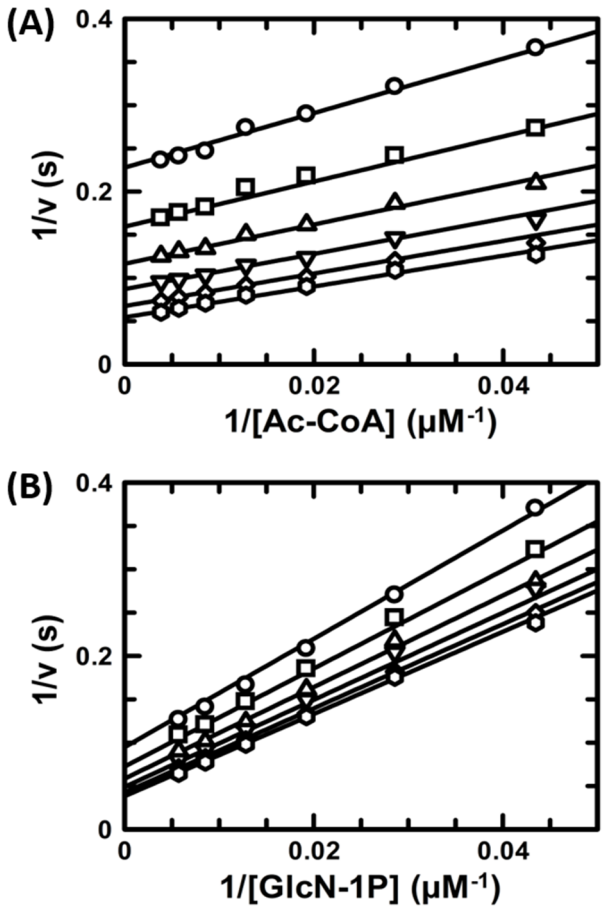


Figure 48. GlmU acetyltransferase initial velocity patterns.

Double-reciprocal plot of GlmU initial rates (A) at varying concentrations of Ac-CoA (> 20 μM) and several fixed concentrations of GlcN-1P: 23 μM (open circles), 35 μM (open squares), 52 μM (open triangles), 78 μM (inverted open triangles), 117 μM (open diamonds) and 175 μM (open hexagonals). (B) at varying concentrations of GlcN-1P (> 20 μM) and several fixed concentrations of Ac-CoA: 23 μM (open circles), 35 μM (open squares), 52 μM (open triangles), 78 μM (inverted open triangles), 117 μM (open diamonds) and 175 μM (open hexagonals). Symbols represent experimental data, and solid lines are fits of the data to Eqn. (4).

The intersecting patterns on double-reciprocal plots (Figure 48) were observed, which are consistent with GlmU catalysed acetyl transfer following a sequential bi-bi kinetic mechanism. Steady state kinetic parameters for GlmU acetyltransferase activity were obtained by fitting these data to Eqn. (4); $K_{m,\ Ac\text{-}CoA}$ = 56.1 ± 1.3 μM, $K_{m,\ GlcN\text{-}1P}$ = 165.6 ± 8.9 μM and k_{cat} = 35.6 ± 1.2 s^{-1}, and are broadly consistent with the previously determined values in Chapter 3.5, which were fit to Eqn. (1). The data shown in Figure 48 were also fitted to an equation describing a Ping-Pong kinetic mechanism, however, this analysis yielded large errors, which indicated that this

mechanism is unlikely to apply to the GlmU acetyltransferase activity. Further information on the kinetic mechanism can be obtained by secondary replots of the slope and intercept values, taken from the data analysis in Figure 48, which are presented in Figure 49.

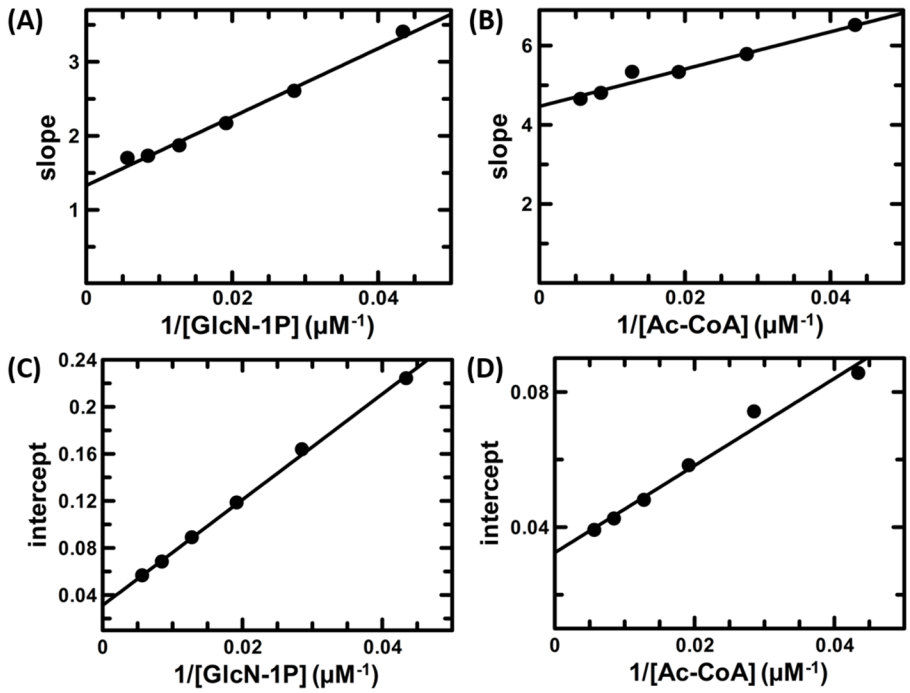


Figure 49. GlmU secondary replots from initial velocity patterns.

(A-B) Secondary replots derived from fixed GlcN-1P concentrations when Ac-CoA is titrated. (A) Replot of slope values from Figure 49 (A). (B) Replot of intercept values from Figure 49 (A). (C-D) Secondary replots derived from fixed Ac-CoA concentrations when GlcN-1P is titrated. (C) Replot of slope values from Figure 49 (B). (D) Replot of intercept values from Figure 49 (B). Symbols represent experimental data, and solid lines are fits of the data to a linear regression.

The re-plotted data presented in Figure 49 demonstrate that a rapid equilibrium ordered mechanism was unlikely, as the slope replot for GlcN-1P did not intersect at 0 (Figure 49 (B)).

Further studies, utilising intact protein ESI-LC-MS, were carried out to investigate if a GlmU acetyl-intermediate is formed upon incubation with Ac-CoA. This study involved the incubation of GlmU with concentrations of either 200 µM or 2 mM Ac-CoA for a minimal time or for 4 hrs, at 25 °C, before injection into the LS-MS. These data were then compared to a GlmU-only control to identify whether any significant intermediate protein species were formed (Figure 50).

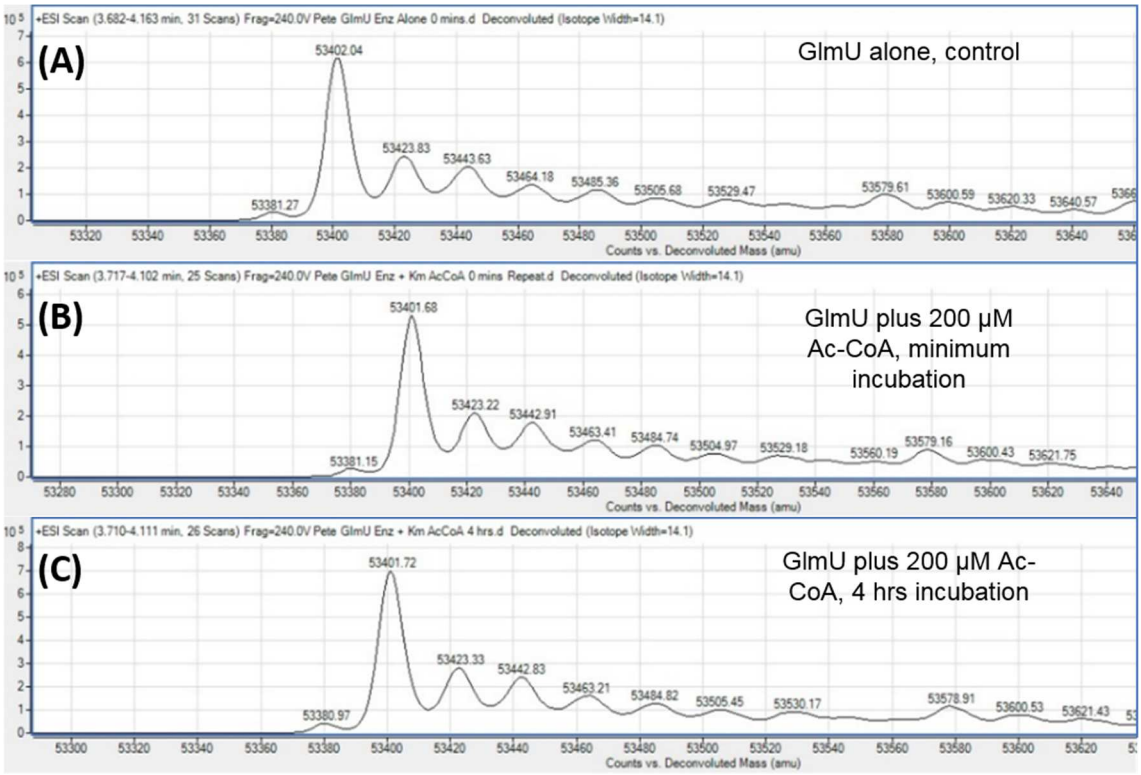


Figure 50. GlmU ESI-LC-MS analysis in the presence of 200 µM Ac-CoA.

Deconvoluted spectra of the intact mass ESI-LC-MS analysis of GlmU to investigate whether an acetyl-intermediate is formed. (A) GlmU alone control, at a concentration of 1.9 µM, injected into the LC-MS after a minimal incubation at 25 °C, (B) GlmU incubated with 200 µM Ac-CoA for a minimal period, at 25 °C, prior to injection into the MS, (C) GlmU incubated with 200 µM Ac-CoA for a 4 hours, at 25 °C, prior to injection into the LC-MS.

The data in Figure 50 demonstrate that no significant acetyl-enzyme intermediate was formed after either the minimal or the 4 hrs incubation and the spectra were identical to the GlmU alone control. The main GlmU species in all three spectra was observed at 53401 Da, which was the predicted molecular weight for the non-modified protein. The data in Figure 51 were very similar and again no acetyl-enzyme intermediate was formed, after incubation with 2 mM Ac-CoA, that was not already present in the GlmU alone control.

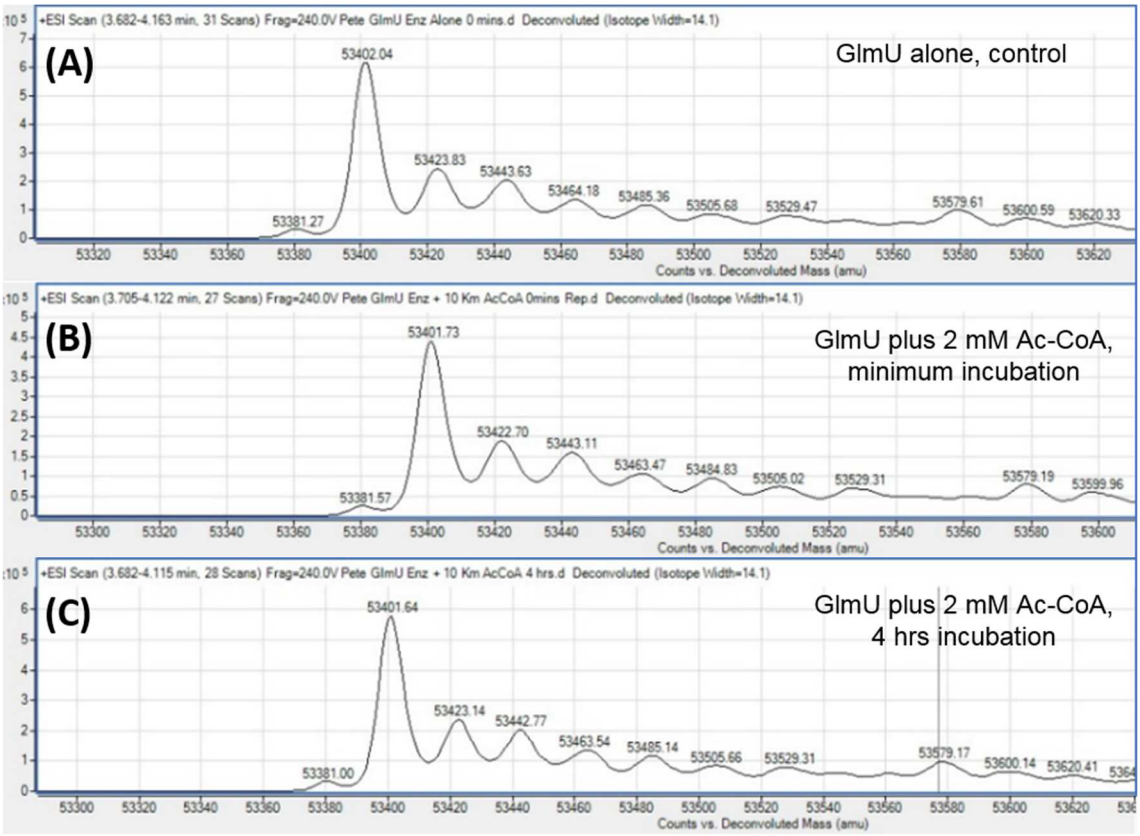


Figure 51. GlmU ESI-LC-MS analysis in the presence of 2 mM Ac-CoA.

Deconvoluted spectra of the intact mass ESI-LC-MS analysis of GlmU to investigate whether an acetyl-intermediate is formed. (A) GlmU alone control, at a concentration of 1.9 µM, injected into the LC-MS after a minimal incubation at 25 °C, (B) GlmU incubated with 2 mM Ac-CoA for a minimal period, at 25 °C, prior to injection into the MS, (C) GlmU incubated with 2 mM Ac-CoA for 4 hours, at 25 °C, prior to injection into the LC-MS.

These ESI-LC-MS investigations confirmed that no covalent acetyl-enzyme intermediates are formed upon incubation, either immediately or after 4 hrs, with two different concentrations of Ac-CoA, which are inconsistent with a likely Ping-Pong acetyl transfer kinetic mechanism.

3.10 Acetyltransferase substrate specificity

To define the substrate specificity of the GlmU acetyltransferase domain, steady-state kinetic parameters were determined from saturation curves, by monitoring CoA-SH release using DTP absorbance, and fitting to Eqn. (1), for a range of GlcN-1P and Ac-CoA analogues (Table 1).

Table 13. Steady-state kinetic parameters for a range of GlmU acetyltransferase substrates [a].

varied substrate	fixed substrate	K_m (μM) [b]	k_{cat} (s^{-1}) [b]	k_{cat}/K_m ($M^{-1}s^{-1}$) [b]	K_d (μM) [b, d]
canonical order					
Ac-CoA	GlcN-1P	250 ± 8	28 ± 1	$1.1 \pm 0.6 \times 10^5$	250 ± 25
GlcN-1P	Ac-CoA	290 ± 31	28 ± 1	$9.7 \pm 1.6 \times 10^4$	ND
UTP	GlcNAc-1P	8 ± 0.8	120 ± 3	$1.5 \pm 0.38 \times 10^7$	ND
GlcNAc-1P	UTP	30 ± 5	100 ± 4	$3.2 \pm 0.8 \times 10^6$	ND
reverse order					
Ac-CoA	UDP-GlcN [c]	40 ± 7	0.27 ± 0.02	$7.5 \pm 2.9 \times 10^3$	250 ± 25
UDP-GlcN	Ac-CoA	9700 ± 600	2.04 ± 0.06	$2.2 \pm 1.7 \times 10^2$	ND
UTP	GlcN-1P	-	-	-	ND
GlcN-1P	UTP	-	-	-	ND
substrate analogues					
AAc-CoA	GlcN-1P	1800 ± 200	0.5 ± 0.27	$2.8 \pm 15 \times 10^2$	ND
GlcN-1P	AAc-CoA [c]	16 ± 6	0.13 ± 0.01	$8.1 \pm 0.5 \times 10^3$	ND
Pro-CoA	GlcN-1P [c]	60 ± 2	0.5 ± 0.01	$8.3 \pm 5 \times 10^3$	500 ± 100
GlcN-1P	Pro-CoA	980 ± 110	0.5 ± 0.01	$5.1 \pm 0.9 \times 10^2$	ND
Suc-CoA	GlcN-1P [c]	1100 ± 100	0.9 ± 0.26	$8.2 \pm 30 \times 10^2$	ND
GlcN-1P	Suc-CoA	-	-	-	ND

[a] Reactions performed at pH 7.5 and 30 °C. [b] Values are mean of at least three experiments ± standard error obtained upon fitting the data to the appropriate equation. [c] Non-saturating concentrations of the fixed substrate used; UDP-GlcN, 800 μM, AAc-CoA, 1 mM, GlcN-1P, 3 mM. All data were fit to Eqn. (1). [d] The binding data was generated utilising DSF by fitting individual titration data to Eqn. (19) and then correlating the change in T_m with titrant concentration and fitting data to Eqn. (20). The following analogues were tested but were not substrates of GlmU acetyltransferase (no activity ≤ 10 mM); Dethio-CoA, Bu-CoA, Eth-CoA, IsoBu-CoA, Mlo-CoA, Cro-CoA (Figure 52), GlcNAc, Glc-1P, GlcN-6P, GalN-1P, galactose-1-phosphate, galactosamine, mannosamine, mannose-1-phosphate, N-acetyl-galactosamine, N-acetyl-mannosamine, glucose-6-phosphate and mannose-6-phosphate.

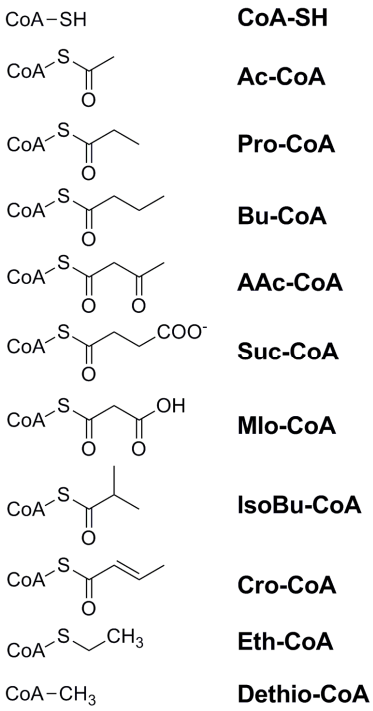


Figure 52. Acyl-CoA and CoA-SH analogues used to investigate GlmU acetyltransferase activity substrate specificity.

CoA-SH analogues or acyl thioesters utilised to investigate the substrate specificity of GlmU acetyltransferase activity.

No acetyltransferase activity was detected, up to a 10 mM concentration, for the following GlcN-1P analogues: galactosamine-1-phosphate, glucosamine, galactosamine, mannosamine, mannose-1-phosphate, mannose-6-phosphate, N-acetyl-glucosamine (GlcNAc), N-acetyl-galactosamine, N-acetyl-mannosamine, glucose-1-phosphate (Glc-1P), glucosamine-6-phosphate (GlcN-6P), glucose 6-phosphate and galactose-6-phosphate. While activity was not observed with most of these potential substrates, inhibition was observed with some of these GlcN-1P analogues (see section 3.11). The only GlcN-1P analogue that was identified as a substrate of the GlmU acetyltransferase activity was UDP-glucosamine (UDP-GlcN, Figure 53), which is further highlighted in Figure 54.

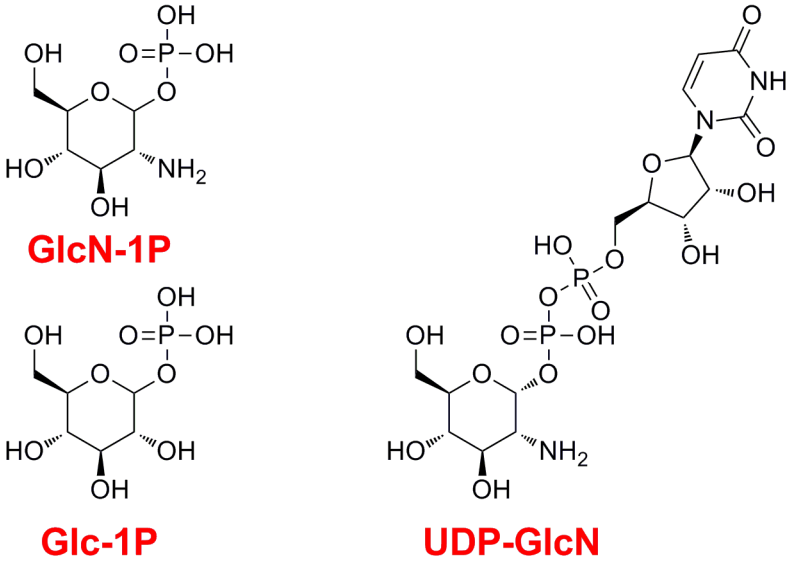


Figure 53. Structure of UDP-GlcN.

Structure of UDP-GlcN. Compound prepared by Martin Rejzek and Rob Field from the Molecules from Nature laboratory, John Innes Centre, Norwich.

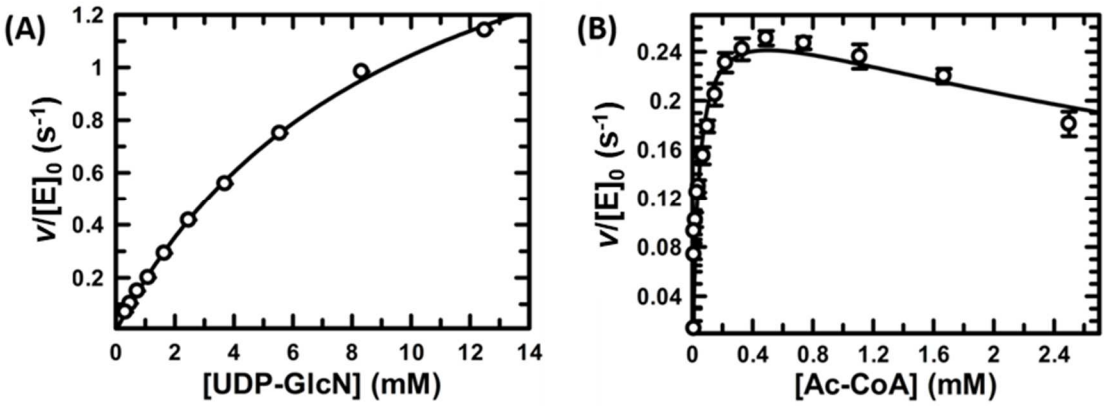
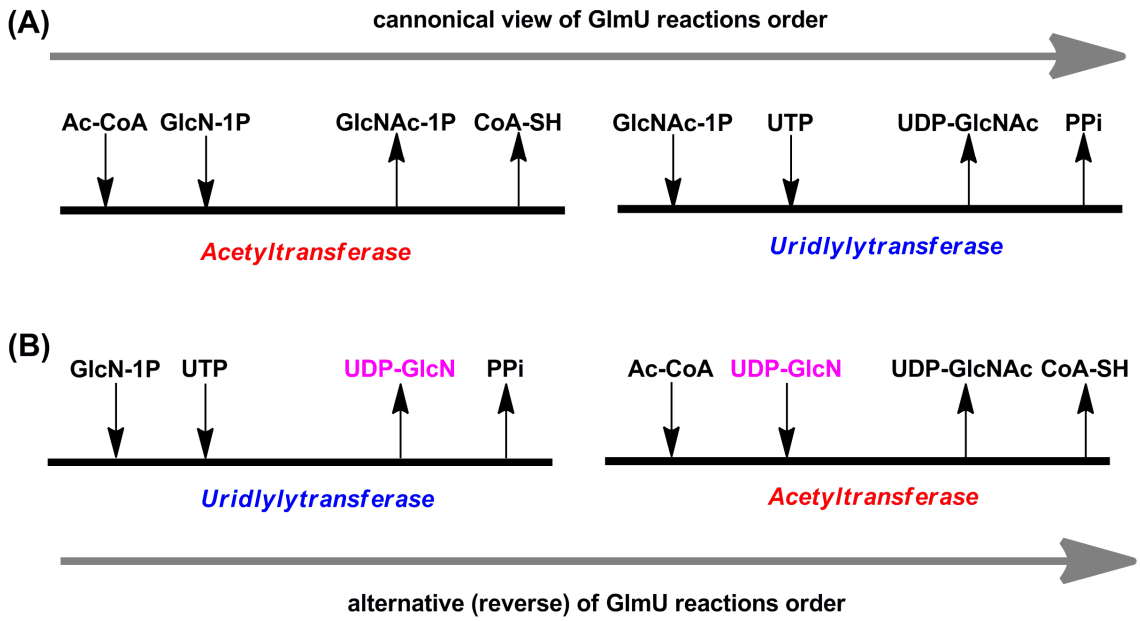


Figure 54. GlmU acetyltransferase steady-state kinetics utilising UDP-GlcN.

Titrations of either UDP-GlcN or Ac-CoA at high concentrations of the second substrate, utilising the DTP absorbance assay to monitor formation of CoA-SH. (A) Titration of UDP-GlcN at a saturating concentration of Ac-CoA. Symbols represent experimental data, and solid lines are fits of the data to Eqn. 1. (B) Titration of Ac-CoA at a non-saturating concentration of UDP-GlcN (800 μM). Symbols represent experimental data, and solid lines are fits of the data to Eqn. (3).

The activity monitored using UDP-GlcN as a second substrate was modest but demonstrates that it is possible to acetylate analogues of GlcN-1P, as well indicating that the catalytic activities of GlmU may not be carried out in the accepted, canonical order. The alternative order of GlmU catalytic activities are compared to the canonical order in Scheme 11.



Scheme 11. GlmU alternative reaction order

The potential alternative or reverse GlmU reaction order proposed in Scheme 11 was not corroborated by the data presented in Table 13, as it was not possible to monitor any uridylyltransferase activity when GlcN-1P, as a substitute for the native substrate GlcNAc-1P, was titrated as a substrate, in the presence of saturating concentrations of UTP. These data do not preclude the existence of an alternative reaction order, but they do suggest that it is not possible for GlmU to catalyse the uridylylation of GlcN-1P. These results do not preclude that other enzymes in *M. tuberculosis* could uridylylate GlcN-1P.

GlmU acyltransferase activity was not detected, up to a 10 mM concentration, for the following acyl-CoAs: ethyl-CoA (Eth-CoA), butyryl-CoA (Bu-CoA), isobutyryl-CoA (IsoBu-CoA), malonyl-CoA (Mlo-CoA), crotonyl-CoA (Cro-CoA), stearoyl-CoA, HMG-CoA, palmitoyl-CoA and glutaryl-CoA. Although these Ac-CoA analogues were not substrates, some were identified as dead-end inhibitors and will be further explored in section 3.11. The Ac-CoA analogues identified as substrates, albeit significantly poorer (k_{cat}/K_m) than the accepted native substrate, were n-propionyl-CoA (Pro-CoA), acetoacetyl-CoA (AAc-CoA), and succinyl-CoA (Suc-CoA), listed in descending order of catalytic efficiency.

The most likely alternative acyl-CoA substrate (highest k_{cat}/K_m) to Ac-CoA, was Pro-CoA and this analogue was subsequently used in initial velocity studies to further

confirm the results in section 3.9 and better understand whether the order of binding to form the ternary complex, was either ordered or random (Figure 55).

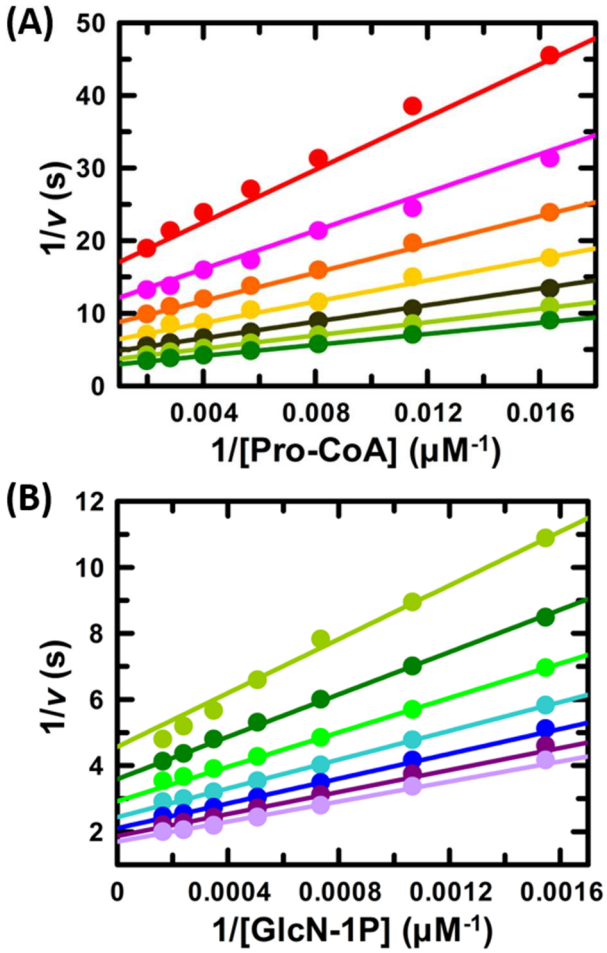


Figure 55. GlmU acetyltransferase initial velocity patterns, using Pro-CoA.

Double-reciprocal plot of GlmU initial rates utilising Pro-CoA as the acyl-CoA substrate. (A) varying concentrations of Pro-CoA (11 – 500 μM) and several fixed concentrations of GlcN-1P: 100 μM (●), 146 μM (●), 212 μM (●), 307 μM (●) 445 μM (●) 646 μM (●) and 936 μM (●). (B) varying concentrations of GlcN-1P (0.1 to 6 mM) and several fixed concentrations of Pro-CoA: 61 μM (●), 87 μM (●), 123 μM (●), 175 μM (●) 248 μM (●) 352 μM (●) and 500 μM (●). Symbols represent experimental data, and solid lines are fits of the data to Eqn. (4).

The patterns on the two double-reciprocal plots in Figure 55, were intersecting, which is again consistent with a sequential bi-bi kinetic mechanism, as was observed when using Ac-CoA. Steady-state kinetic parameters for GlmU acetyltransferase activity using Pro-CoA were obtained by fitting these data to Eqn. (4); $K_{m,\ Pro\text{-}CoA}$ = 172 ± 14.5 μM, $K_{m,\ GlcN\text{-}1P}$ = 1104 ± 80.2 μM and k_{cat} = 0.85 ± 0.04 s^{-1}, which are broadly consistent with the previously determined values in Table 13, which were fit to Eqn. (1). Additional information on the kinetic mechanism can be attained by secondary

replots of the slope and intercept values, taken from the data analysis in Figure 55, which are presented in Figure 56.

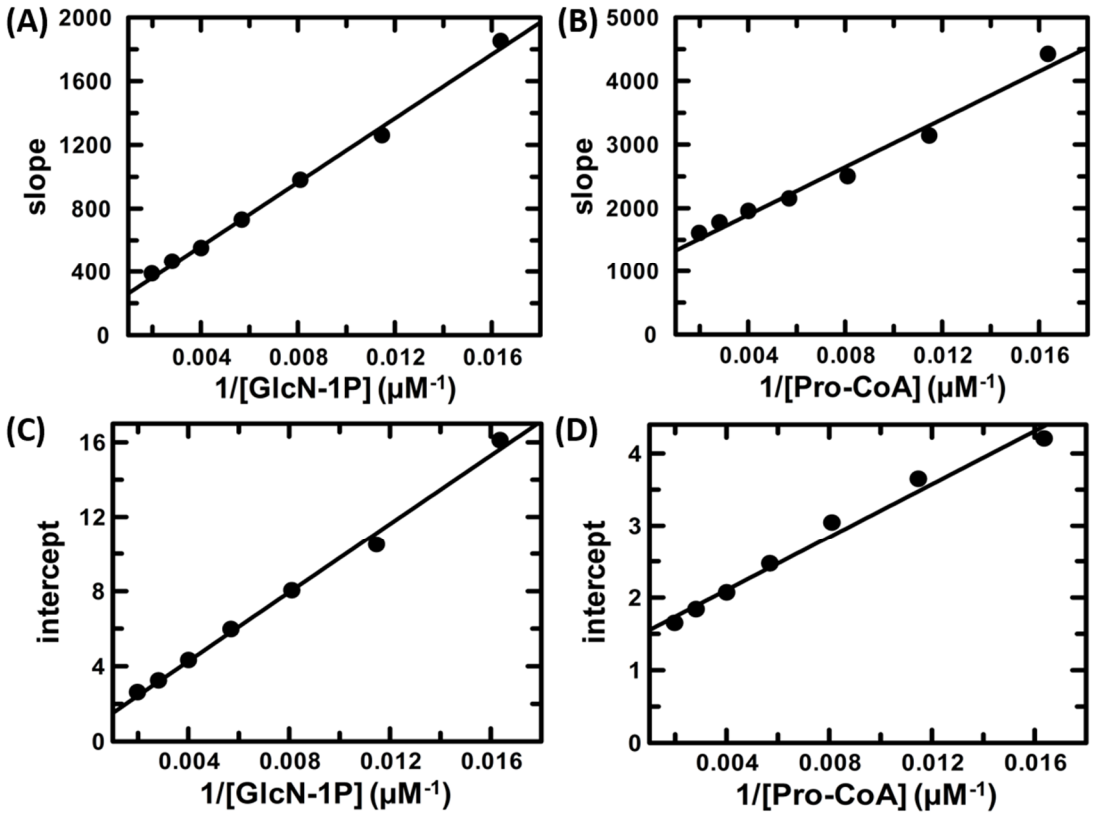


Figure 56. GlmU secondary replots from Pro-CoA initial velocity patterns.

(A-B) Secondary replots derived from fixed GlcN-1P concentrations when Pro-CoA is titrated. (A) Replot of slope values from Figure 55Figure 48 (A). (B) Replot of intercept values from Figure 55 (A). (C-D) Secondary replots derived from fixed Pro-CoA concentrations when GlcN-1P is titrated. (C) Replot of slope values from Figure 55 (B). (D) Replot of intercept values from Figure 55 (B). Symbols represent experimental data, and solid lines are fits of the data to a linear regression.

The data presented in Figure 56, which confirm the observations in Figure 49, demonstrate that a rapid equilibrium ordered mechanism was unlikely, as the slope replot for GlcN-1P did not intersect at 0 (Figure 56 B).

Finally, to provide further information about the order of substrate binding to form the ternary complex prior to acetyl transfer, initial binding studies were carried out to elucidate whether the canonical substrates were able to bind to the apoenzyme. DSF thermal stability studies, utilising SYPRO Orange, were undertaken using a fixed concentration of GlmU that had been pre-incubated with titrations of either of the acetyltransferase substrates (Figure 57).

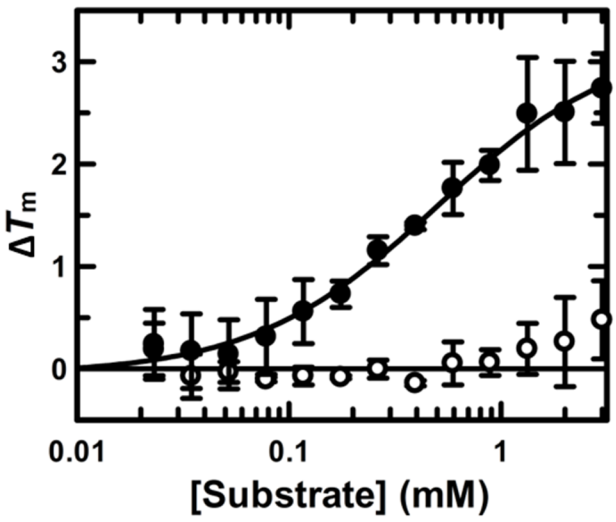


Figure 57. DSF studies investigating binding of acetyltransferase substrates to GlmU.

Titrations of either Ac-CoA (●) or GlcN-1P (○) in the presence of 1.5 μM GlmU and 10x SYPRO Orange. For each condition, the data are a mean of 3 discrete replicates. The data for each condition investigated were fitted to Eqn. (19), to determine the T_m. No substrate, GlmU melting curves were carried out as controls. Subsequently, the change in T_m (ΔT_m) was determined for each test condition and for Ac-CoA fitted to Eqn. (20) to determine the apparent K_d for this binding event.

The data in Figure 57 demonstrates that, using DSF, it was possible to monitor a change in GlmU T_m , from an unliganded T_m = 49.8 ± 0.05 °C , to T_m > 52.4 °C in the presence of Ac-CoA concentrations greater than 1 mM. There was a clear GlmU T_m stabilisation in the presence of an Ac-CoA titration from 2 mM, which when fit to Eqn. (20) yielded a K_d = 390 ± 140 μM. A stabilisation of the GlmU T_m of 0.5 °C at 2 mM GlcN-1P was observed, however, there is no clear concentration dependence and the data could not be fit to Eqn. (20). These data indicate that Ac-CoA can bind to and stabilise free GlmU, with a K_d that is broadly in line with $K_{m,\ Ac\text{-}CoA}$ (Figure 39). In contrast, it was not possible to stabilise free GlmU with GlcN-1P in the same concentration range, which is indicative of only modest binding of the substrate to the unliganded enzyme.

3.11 Product and dead-end inhibition patterns

In the previous section, several acyl-CoA compounds, along with two GlcN-1P analogues, were identified to be inhibitors of the GlmU acetyltransferase activity,

when assayed in the presence of the canonical substrates, which were used initially at concentrations equivalent to the determined K_m values (Table 2).

Table 14. Substrate analogue and product inhibition parameters for GlmU acetyltransferase activity [a].

Inhibitor	IC_{50} (mM) [b, c]	n^H [b, c]	K_d (mM) [c, e]
CoA-SH	0.3 ± 0.01	0.9 ± 0.1	0.7± 0.13
Eth-CoA [d]	ND*		ND
Bu-CoA	1 ± 0.03	1.2 ± 0.1	3.8 ± 0.8
IsoBu-CoA	2 ± 0.21	1.5 ± 0.2	ND
Mlo-CoA	2 ± 0.59	1.1 ± 0.6	ND
Cro-CoA	1 ± 0.15	1.2 ± 0.1	ND
GlcNAc [d]	ND*	ND	ND
Glc-1P	5 ± 0.65	1.0 ± 0.1	ND
GlcN-6P	10 ± 0.89	1.1 ± 0.1	ND
GalN-1P	120 ± 30	1.1 ± 0.1	ND

[a] Reactions performed at pH 7.5 and 30 °C. [b] Reactions were carried out using concentrations equivalent to acetyltransferase K_m for Ac-CoA and GlcN-1P, competitors titrated from 20 μM. [c] Values are mean of at least three experiments ± standard error obtained upon fitting the data to the appropriate equation. The inhibition data was to Eqn. (8) [d] Ligands that show does dependent inhibition but at the highest tested concentration do not fully inhibit GlmU acetyltransferase activity. [e] The binding data was generated utilising DSF by fitting individual titration data to Eqn. (19) and then correlating the change in T_m with titrant concentration and fitting data to Eqn. (20). The following analogues were tested but were not inhibitors of GlmU acetyltransferase (no inhibition ≤ 10 mM); GlcNAc-1P, Dethio-CoA, galactose-1-phosphate, galactosamine, mannosamine, mannose-1-phosphate, N-acetyl-galactosamine, N-acetyl-mannosamine, glucose-6-phosphate and mannose-6-phosphate.

In Table 14, it is apparent that CoA-SH was identified as an inhibitor, which, using the previously described thiol detection assays, would not be possible due to reaction of either DTP or DTNB with the free thiol of the product. To investigate product inhibition, it was necessary to prepare samples in an alternate manner or utilise alternative assays to monitor the effect of titrations of CoA-SH on the acetyltransferase reaction (Figure 58).

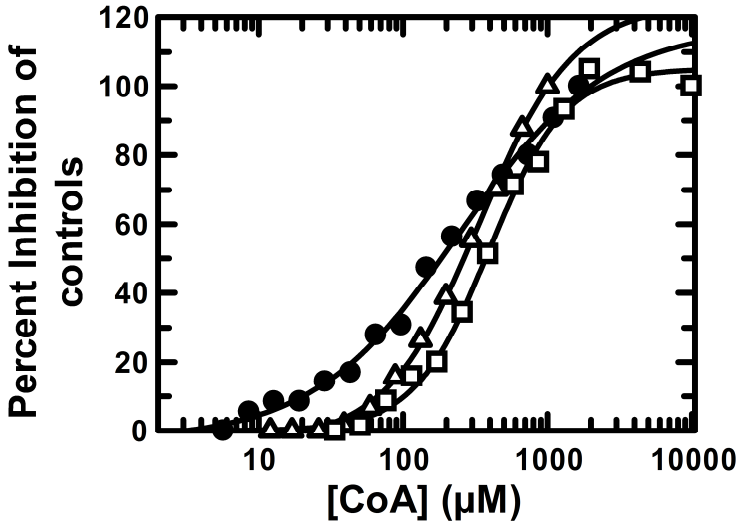


Figure 58. GlmU CoA-SH or CoA-NEM product inhibition.

CoA was titrated against fixed concentrations, equivalent to the respective K_m values for Ac-CoA and GlcN-1P, and GlmU acetyltransferase was monitored using three differing methodologies. Solid circles (●) were data generated using a fluorescent, enzyme-coupled, phosphate detection methodology to monitor the effect of CoA-SH on phosphate accumulation. Open triangles (Δ) were data generated using a mass spectrometry approach to monitor the effect of CoA-SH on the conversion of GlcN-1P to GlcNAc-1P. Open squares (□) were data generated using CoA-SH that was reacted with an equimolar concentration of NEM, prior to monitoring GlmU mediated CoA-SH accumulation using DTP. All data were normalised to the separate assay controls for no and full inhibition of GlmU acetyltransferase activity. Symbols represent experimental data, and solid lines are fits of the data to Eqn. (8).

The data in Figure 58 when fitted to Eqn. (8) allowed determination of the following inhibition values: IC_{50} = 300 ± 25 μM (Figure 58, ●), using an enzyme coupled fluorescent assay to monitor uridylyltransferase activity (Section 2.2.10); IC_{50} = 250 ± 40 μM (Figure 58, Δ), using a Rapidfire-MS (RF-MS) assay to monitor acetylation of GlcN-1P to form GlcNAc-1P (Section 2.2.11) and IC_{50} = 390± 30 μM (Figure 58, □), using the DTP thiol assay and a modified form of CoA-SH reacted with NEM. These IC_{50} values are broadly in agreement, within the range from 250 to 390 μM, and confirm that the acetyltransferase reaction is inhibited by the product CoA-SH. These data were further corroborated by DSF binding studies, which were carried out by titrating CoA-SH or Ac-CoA before incubation with 2 μM GlmU and subsequent thermal denaturation in the presence of SYPRO Orange (Figure 59).

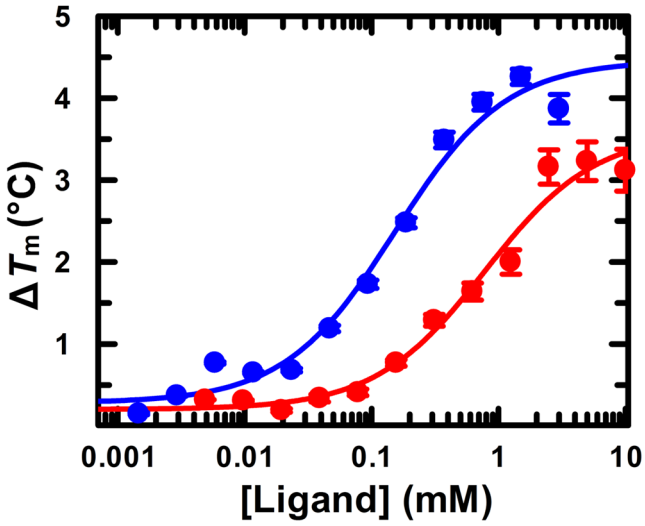


Figure 59. DSF studies investigating binding of CoA-SH to GlmU.

Titrations of either Ac-CoA (●) or CoA-SH (●) in the presence of 2 μM GlmU and 10x SYPRO Orange. For each condition, the data are a mean of 2 discrete replicates. The data for each condition investigated were fit to Eqn. (19), to determine the T_m. No substrate, GlmU melting curves were carried out as controls (data not shown). Subsequently, the change in T_m (ΔT_m) was determined for each test condition and the titration data for both Ac-CoA and CoA-SH were fitted to Eqn. (20) to determine the apparent K_d for these binding events.

The data in Figure 59 further demonstrates that, using DSF, it was possible to monitor an increase in GlmU T_m in the presence of high concentrations of either CoA-SH or Ac-CoA. There was a clear GlmU T_m stabilisation in the presence of a CoA-SH titration of CoA-SH (from 10 mM), which when fitted to Eqn. (20) yielded an apparent K_d = 780 ± 170 μM. The control titration of Ac-CoA, when fitted to Eqn. (20) yielded an apparent K_d = 150 ± 30 μM, which is two-fold more potent than previously determined for this interaction (Figure 57). These data demonstrate that, in addition to Ac-CoA, CoA-SH can bind to and stabilise free GlmU, with a K_d that is largely in agreement with the IC_{50} values determined for product inhibition in Figure 58.

To investigate the order of substrate binding to, and product release from, the GlmU acetyltransferase domain, product inhibition pattern studies were carried out, the results of which are detailed in Table 15.

Table 15. Product inhibition patterns and inhibition constants for GlmU acetyltransferase activity[a].

varied substrate	inhibitor	fixed substrate	inhibition pattern [c]	K_{is} (mM)	K_{ii} (mM)
product inhibition					
GlcN-1P	CoA-SH	10×K_m Ac-CoA	NC	0.21 ± 0.086	0.8 ± 0.09
Ac-CoA	CoA-SH	10×K_m GlcN-1P	C	0.15 ± 0.015	
GlcN-1P	GlcNAc-1P	0.5 to 10×K_m Ac-CoA	NI		
Ac-CoA	GlcNAc-1P	0.5 to 10×K_m GlcN-1P	NI		
dead-end inhibition					
GlcN-1P	Glc-1P	10×K_m Ac-CoA	C	4.1 ± 0.6	
Ac-CoA	Glc-1P	0.5 to 10×K_m GlcN-1P [b]	UC		11.56 ± 0.9
GlcN-1P	Bu-CoA	10×K_m Ac-CoA	C	1.2 ± 0.2	
Ac-CoA	Bu-CoA	10×K_m GlcN-1P	C	0.35 ± 0.03	

[a] Reactions performed at pH 7.5 and 30 °C. [b] Range of GlcN-1P concentrations used from 0.5 to 10 × K_m. [c] NC, non-competitive, fit to Eqn. (6); C, competitive, fit to Eqn. (5); UC, uncompetitive, fit to Eqn. (7); NI, no inhibition.

Overall, the results detailed in Table 15 were consistent with the acetyltransferase activity following an ordered sequential mechanism, which is primarily supported by the patterns determined using CoA-SH. This product was found to competitive versus Ac-CoA, which confirms that both these co-factors compete for the same binding site on the free enzyme. This competitive pattern suggests that Ac-CoA is the first substrate to bind and CoA-SH is the last product to dissociate from the acetyltransferase domain. The noncompetitive pattern determined for CoA-SH versus GlcN-1P, further suggests a sequential ordered mechanism is likely, as this mode of inhibition is characterised by the inhibitor binding to the free enzyme and the enzyme-substrate binary complex. Importantly, for noncompetitive inhibition the inhibitor and substrate do not compete for binding to the free enzyme, which suggests that CoA-SH and GlcN-1P binding sites are spatially distinct. These findings would support a mechanism whereby Ac-CoA binds first to the acetyltransferase domain followed by GlcN-1P. No inhibition was observed when GlcNAc-1P was tested versus either Ac-CoA or GlcN-1P, at saturating and several sub-saturating concentrations of the respective co-substrate. These results suggest

an equilibrium ordered kinetic mechanism, with GlcNAc-1P being the first product to dissociate from the acetyltransferase ternary complex followed by CoA-SH.

To further confirm that the GlmU acetyltransferase activity follows an equilibrium ordered kinetic mechanism, dead-end inhibition experiments were carried out using analogues that were identified as inhibitors, see Table 14. The dead-end inhibitors selected for these studies were Bu-CoA and Glc-1P, which were initially validated for mode of inhibition by carrying out mutual exclusivity studies, which involved titrating Bu-CoA at several fixed concentrations at Glc-1P, in the presence of concentrations equivalent to the K_m values of Ac-CoA and GlcN-1P and monitoring CoA-SH release by DTP absorbance change (Figure 60).

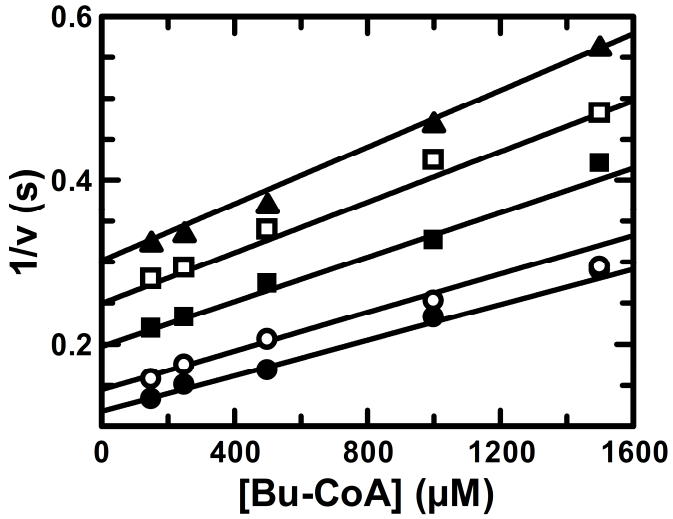


Figure 60. GlmU acetyltransferase mutual exclusivity studies using the dead-end inhibitors Bu-CoA and Glc-1P.

A Dixon plot, also known as a Yonetani-Theorell plot, which was arranged by plotting the initial velocities for titrations of Bu-CoA at several fixed concentrations of Glc-1P: 2.5 mM (circles), 5 mM (open circles), 10 mM (squares), 15 mM (open squares) and 20 mM (triangles). Symbols represent experimental data, and solid lines are fits of the data to Eqn. (22).

The data in Figure 60 when fitted to Eqn. (22) allows determination of two inhibition constants, K_i and K_j, for the titrants; Bu-CoA, K_i = 950 ± 85 μM and Glc-1P, K_j = 8.9 ± 0.6 mM, which are in agreement with the IC_{50} values presented in Table 14. In addition to determination of the inhibition constants, an α value is calculated, which is a measure of co-operativity of the two inhibitors. Furthermore, the patterns of the lines on the Dixon plot are diagnostic of whether the inhibitors compete for the same binding site on GlmU. The calculated α = 3, which is indicative of the two inhibitors

antagonising each other's binding, along with the lines on the plot in Figure 60 that eventually intersect below the x-axis, indicate that Bu-CoA and Glu-1P, bind in a mutually exclusive fashion, competing with one another for the same GlmU form.

The dead-end inhibition pattern results presented in Table 15 taken together, agreed with GlmU acetyltransferase following an ordered, sequential mechanism. Competitive inhibition patterns were obtained for both Bu-CoA versus Ac-CoA and GlcN-1P, which are indicative of Bu-CoA competing for binding to the free enzyme and preventing GlcN-1P from binding, likely by steric hindrance. In addition, a competitive inhibition pattern was observed for Glc-1P versus GlcN-1P, which was expected as the structures differ only by the presence of an amine group at the C2 position of the hexose ring. The determination of an uncompetitive inhibition pattern for Glc-1P versus Ac-CoA indicates that Glc-1P binds to either the E:Ac-CoA or E:CoA-SH complex, and not to the free enzyme form. The Glc-1P versus Ac-CoA uncompetitive inhibition pattern was further investigated by repeating the experiment, while reducing the co-substrate, GlcN-1P, to sub-saturating concentrations (Figure 61).

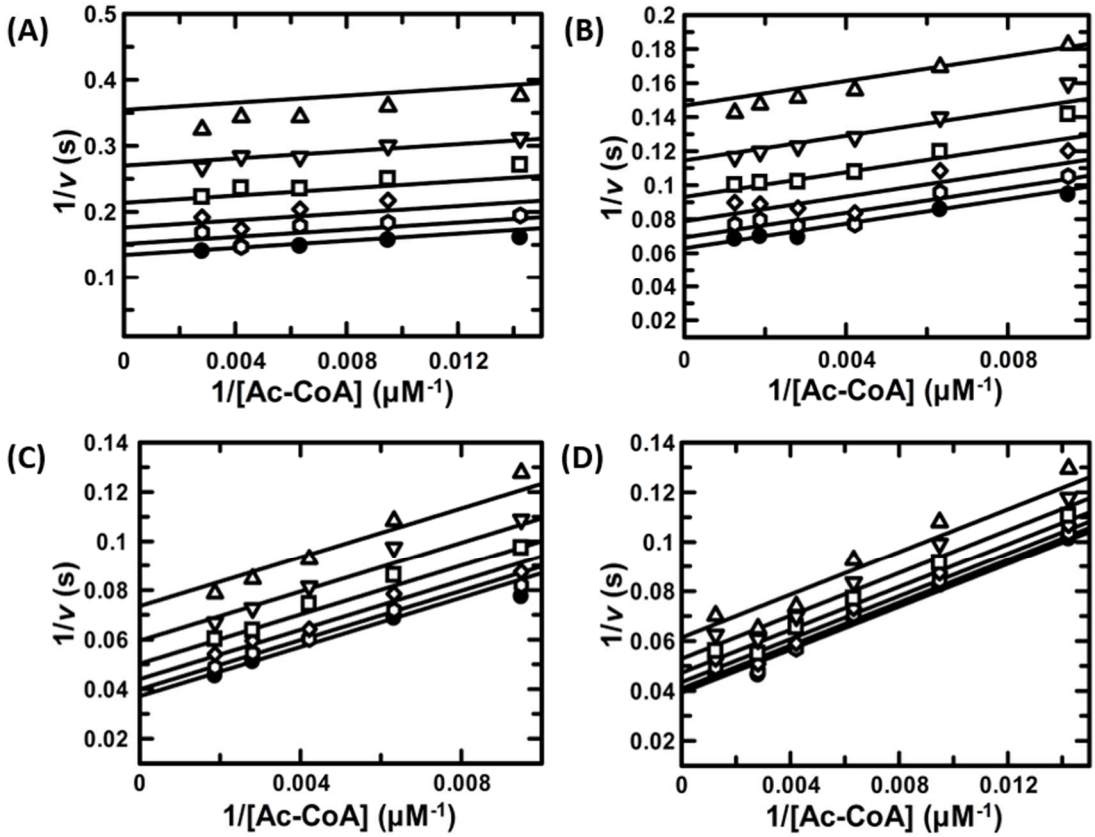


Figure 61. Uncompetitive inhibitor patterns for GlmU acetyltransferase

(A-D) Titrations of Ac-CoA at several fixed concentrations of Glc-1P: 20 mM (triangles), 13.3 mM (inverted triangles), 8.8 mM (squares), 6 mM (diamonds), 4 mM (hexagonals) and 2.6 mM (filled circles). (A) Experiment carried out in the presence of 100 µM of GlcN-1P, (B) Experiment carried out in the presence of 300 µM of GlcN-1P, (C) Experiment carried out in the presence of 600 µM of GlcN-1P and (D) Experiment carried out in the presence of 1000 µM of GlcN-1P. Symbols represent experimental data, and solid lines are fits of the data to Eqn. (7). Fitting to Eqn. (7) allows determination of a K_{ii}, which for the experimental data shown here were: (A) K_{ii} = 8 ± 0.6 mM; (B) K_{ii} = 10 ± 0.7 mM; (C) K_{ii} = 15 ± 0.9 mM; (D) K_{ii} = 28 ± 0.4 mM.

As the concentration of GlcN-1P is reduced, the inhibitor pattern for Glc-1P versus Ac-CoA remains uncompetitive and the determined K_{ii} reduces from 28 ± 0.4 mM, for the 1000 µM GlcN-1P test condition, to 8 ± 0.6 mM, for the 100 µM experiment. These uncompetitive inhibitor patterns confirm that Glc-1P binds to either the E:Ac-CoA or E:CoA-SH complex and the increase in $K_{ii,}$ as the concentration of GlcN-1P was increased was indicative of the competitive inhibition between Glc-1P and GlcN-1P.

To further confirm the formation of an ordered ternary complex, with Ac-CoA binding to the free enzyme, followed by GlcN-1P, binding studies using DSF were carried out. These studies were a continuation of the DSF binding experiments presented in Figure 57 and Figure 59 and involved incubation of GlmU with saturating concentrations of either Ac-CoA or CoA-SH, followed by addition of titrations of Glc-1P to either the ES or EP complex, in the presence of SYPRO Orange, prior to exposure of each test condition to a temperature gradient (Figure 62).

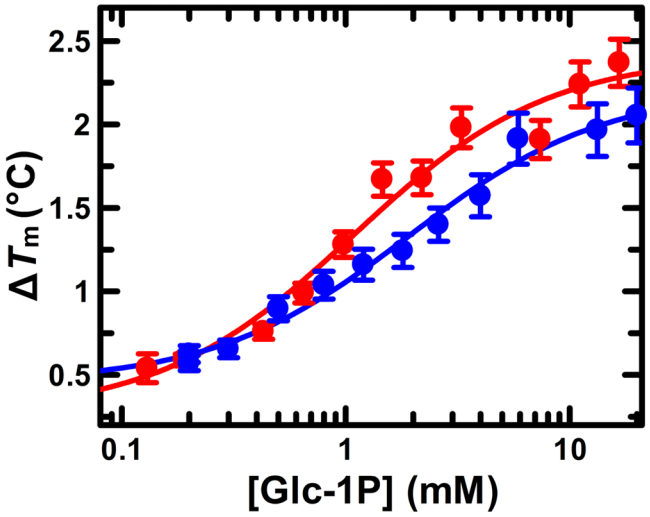


Figure 62. DSF studies investigating GlmU acetyltransferase ternary complex formation.

Titrations of Glc-1P in the presence of 1.5 or 2 μM GlmU pre-incubated with either Ac-CoA (●) or CoA-SH (●), respectively and 10x SYPRO Orange. For each condition, the data are a mean of either 2 or 4 discrete replicates. The data for each condition investigated were fit to Eqn. (19), to determine the T_m. GlmU in the presence of either 1 mM Ac-CoA or 10 mM CoA-SH melting curves were carried out as controls (data not shown). Subsequently, the change in T_m (ΔT_m) was determined for each test condition and the Glc-1P titration data for both the ES or EP complexes were fitted to Eqn. (20) to determine the apparent K_d for these binding events.

In the presence of a saturating concentration of either Ac-CoA or CoA-SH, it was possible to observe an additional, Glc-1P concentration-dependent stabilisation of GlmU T_m. As was observed with GlcN-1P, in Figure 57, using DSF, it was not possible to monitor any stabilisation above background after incubation of the apoenzyme with a titration of Glc-1P (data not shown). The Glc-1P concentration dependent stabilisation of GlmU T_m, when fitted to Eqn. (20), which yielded an apparent $K_{d,\,ES}$ = 1.9 ± 0.4 mM and $K_{d,\,EP}$ = 1.1 ± 0.3 mM, which are like the Glc-1P IC_{50} values previously determined. These DSF results demonstrate that GlmU T_m can only be stabilised by Glc-1P when the enzyme is already in complex with either Ac-CoA or CoA-SH. These data suggest that Glc-1P is unlikely to bind first to the free enzyme and can only bind in an ordered sequence to form GlmU ternary complexes.

Additional DSF studies were carried out to investigate binding of substrate analogues and dead-end inhibitors, using the same experimental approach that had been utilised for the aforementioned Ac-CoA, CoA-SH and Glc-1P studies (Table 16).

Table 16. Dissociation Constants for GlmU with different ligands [a].

ligand	second ligand [b]	K_d (μM)
Ac-CoA	-	250 ± 25
GlcN-1P	-	N.D. [c]
Glc-1P	-	N.D. [c]
Glc-1P	Ac-CoA	1700 ± 400
Bu-CoA	-	3800 ± 800
GlcN-1P	Bu-CoA	N.D.
Pro-CoA	-	500 ± 100
CoA-SH	-	700 ± 130
Glc-1P	CoA-SH	1100 ± 300

[a] Experiments performed at pH 7.5. [b] Co-substrate concentrations used: Ac-CoA, 2 mM; CoA-SH, 10 mM. [c] Highest concentration tested was 3000 μM. The binding data was generated utilising DSF by fitting individual titration data to Eqn. (19) and then correlating the change in T_m with titrant concentration and fitting data to Eqn. (20).

The data presented in Table 16 confirm that either CoA-SH or acyl-CoA can bind to and stabilise the free enzyme T_m whereas GlcN-1P and analogues cannot, which was indicative of sequential ordered ternary complex formation.

Finally, to further confirm the kinetic mechanism of GlmU acetyltransferase, inhibitor pattern studies were carried out using Pro-CoA, as opposed to Ac-CoA as the acyl-CoA substrate. Initially, to confirm inhibition of GlmU propionyl transfer by either CoA-SH or the previously identified dead-end analogues, the potential inhibitors were titrated in the presence of fixed concentrations of GlmU, Pro-CoA and GlcN-1P and CoA-SH release monitored using the DTP absorbance assay (Figure 63).

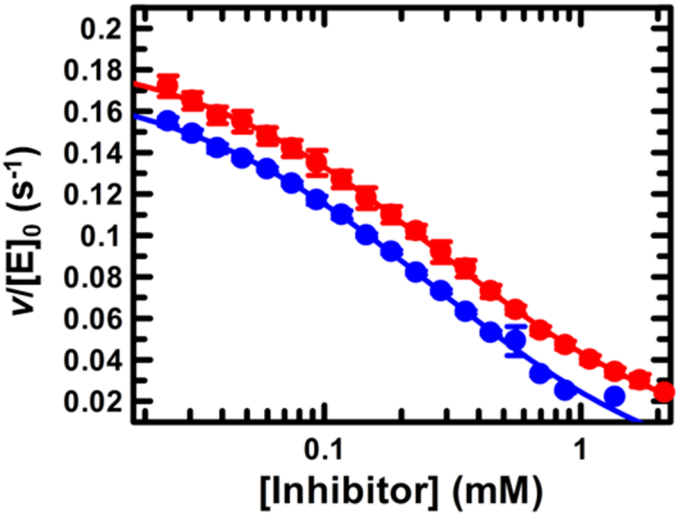


Figure 63. Inhibition of GlmU propionyltransferase activity.

Potential inhibitors were titrated against fixed concentrations, equivalent to the respective K_m values for Ac-CoA and GlcN-1P, and GlmU propionyltransferase was monitored using the DTP absorbance assay. Bu-CoA (●) and CoA-SH, reacted with NEM (●) were titrated from 3.3 mM, while Glc-1P (●) was titrated from 33.3 mM before incubation with GlmU, prior to initiation of propionyl transfer with 100 µM Pro-CoA and 1 mM GlcN-1P. Symbols represent experimental data, and solid lines are fits of the data to Eqn. (8).

Product and dead-end analogues, previously characterised against the acetyltransferase activity, were evaluated as potential inhibitors against GlmU catalysed propionyl transfer. Fitting the inhibitor titration data in Figure 63 to Eqn. (8) allowed determination of the following values; $IC_{50,\ Bu\text{-}CoA}$ = 280 ± 10 µM and $IC_{50,\ CoA\text{-}NEM}$ = 264 ± 21 µM. The measured propionyl transfer IC_{50} for CoA-NEM was the same for the acetyltransferase value (Table 14, Figure 58), which confirmed that inhibition is due to direct competition between the product and substrate for binding to the GlmU acetyltransferase domain, regardless of the substrate utilised. These observations are further confirmed by the inhibition data generated utilising Bu-CoA as the dead-end analogue. The propionyltransferase $IC_{50,\ Bu\text{-}CoA}$ was 3-fold more potent than the acetyltransferase value, which indicated that the extended length of acyl-chain is more inhibitory for propionyl transfer. The inhibition of GlmU propionyl transfer by Glc-1P is weakened by 7-fold, to an IC_{50} = 42 ± 20 mM when compared to the IC_{50} for acetyltransferase activity (data not shown). These weakened IC_{50} values are likely due to the increased length of the Pro-CoA acyl-chain, which is analogous to the 5-fold reduction in $K_{m,\ GlcN\text{-}1P}$ for propionyltransferase activity.

Inhibitor pattern studies were carried out for GlmU catalysed propionyl transfer using Bu-CoA as a dead-end competitor to confirm the observations in Figure 63 ().

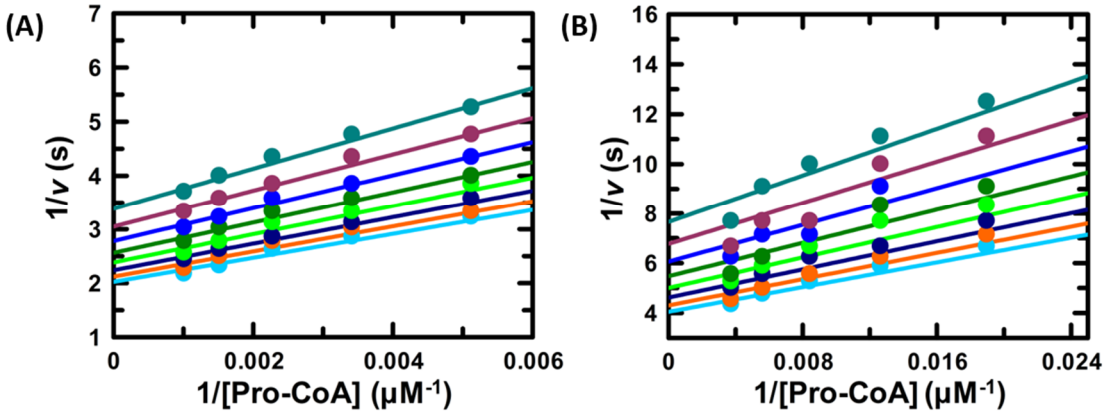


Figure 64. Bu-CoA inhibitor patterns for GlmU propionyltransferase.

Titrations of Pro-CoA at several fixed concentrations of Bu-CoA: 660 µM (●), 540 µM (●), 440 µM (●), 360 µM (●), 290 µM (●), 235 µM (●), 190 µM (●), and 155 µM (●). (A) Experiment carried out in the presence of 1 mM of GlcN-1P and (B) Experiment carried out in the presence of 5 mM of GlcN-1P. Symbols represent experimental data, and solid lines are fits of the data to Eqn.(6).

The model that best fit the experimental data presented in Figure 64 was the noncompetitive inhibition equation Eqn. (6). It was not possible to use saturating concentrations of GlcN-1P, due to reagent limitations, but the inhibitor pattern was carried out at 1 mM and 5 mM (concentrations equivalent to K_m and 5-fold K_m), which did not change the model that best fit the data. The Bu-CoA K_{ii} determined for the 1 mM GlcN-1P conditions was 598 ± 23 µM and for the 5 mM experiment, the Bu-CoA K_{ii} was 412 ± 28 µM. The apparent $K_{m,\ Pro\text{-}CoA}$ for these experiments were 110 ± 3.9 µM and 31 ± 1.6 µM, respectively for the 1 mM and 5 mM GlcN-1P test conditions. Noncompetitive inhibition patterns suggest that Bu-CoA does not bind to the free enzyme and when it binds to the ES complex it will be at a site distal to the active site. These observations suggest that the acyl-CoA and amino sugar binding sites are distinct and binding of Bu-CoA, due to the longer acyl-chain than Ac-CoA, overlaps into the second substrate binding site and impedes GlcN-1P binding.

Collectively, the findings from the inhibitor studies, demonstrate that sequential, ordered ternary complex formation is the most likely kinetic mechanism for the GlmU acetyltransferase activity. To further understand this kinetic mechanism, X-ray

crystallography studies were carried out to capture the Glc-1P, dead-end inhibited ternary complex.

3.12 Structure of acetyltransferase ternary complex

To confirm and rationalise inhibition of GlmU acetyltransferase activity by the dead-end inhibitor, Glc-1P, as well as further understand the catalytic ternary complex, GlmU protein crystals were soaked with 10 mM Ac-CoA and Glc-1P. This approach was successful and allowed determination of the crystal structure of the E:Ac-CoA:Glc-1P complex, to a resolution of 2.26 Å, uniquely, without the addition of the uridylyltransferase product UDP-GlcNAc, which had previously been postulated to be critical for stabilisation of the E:Ac-CoA complex. The data collection and refinement statistics for the GlmU Glc-1P:Ac-CoA complex crystals are detailed in Table 17. These experiments were carried out and analysed by Stephane Mouilleron, Structural Biology Platform, The Francis Crick Institute.

Table 17. Data collection and refinement statistics for GlmU Glc-1P:Ac-CoA.

data collection	
Resolution range	45.71 - 2.26 (2.34 - 2.26)
Space group	H 3 2
Unit cell	109.7 109.7 364.8 90 90 120
Total reflections	260 819 (25 211)
Unique reflections	40 178 (3 861)
Multiplicity	6.5 (6.3)
Completeness (%)	98.3 (96.6)
Mean I/sigma (I)	9.6 (1.1)
Wilson B-factor	49.2
R-merge	0.07 (1.07)
R-pim	0.023 (0.46)
CC1/2	0.995 (0.934)
refinement	
Unique Reflections	39 542 (3 859)
R-work	0.23 (0.48)
R-free	0.26 (0.50)
Number of atoms	3670

macromolecules	3483
ligands	94
solvent	93
RMS (bonds)	0.002
RMS (angles)	0.60
Ramachandran favored (%)	97.5
Ramachandran allowed (%)	2.54
Ramachandran outliers (%)	0.0
Average B-factor	67.0
macromolecules	67.1
ligands	75.3
solvent	57.2

The overall molecular structure of the E:Ac-CoA:Glc-1P complex is similar to that of the previously reported ternary complex of *M. tuberculosis* GlmU E:CoA-SH:GlcN-1P and the complexes superpose well with a root-mean-square deviation of 0.46 Å over 414 Cα. There are however differences between the two structures, notably the orientation of the β-mercaptoethylamine group of either CoA-SH or Ac-CoA. The structure of the E:Ac-CoA:Glc-1P complex displayed the conserved GlmU two-domain architecture, organised into a trimer formed around a threefold crystallographic symmetry (Figure 65).

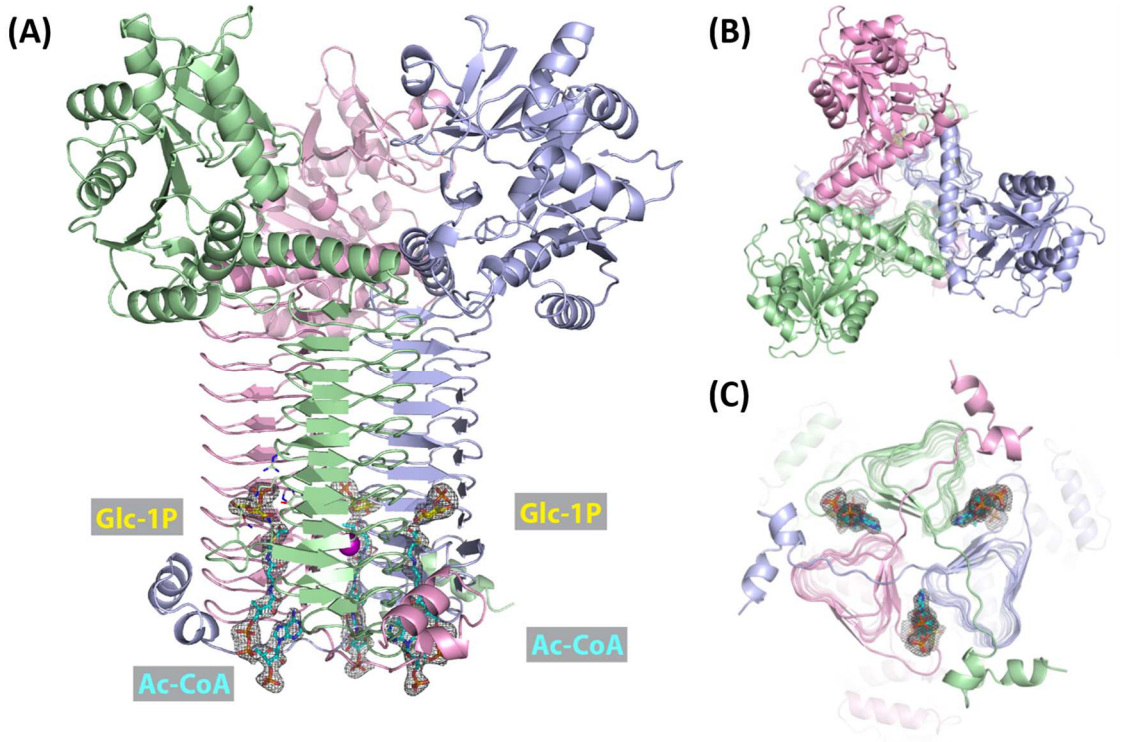


Figure 65. Overview of the GlmU trimeric assembly in complex with Glc-1P and Ac-CoA.

These experiments were carried out and analysed by Stephane Mouilleron, Structural Biology Platform, The Francis Crick Institute. The figure depicts the GlmU quaternary structure in complex with Glc-1P and Ac-CoA. Each monomer is displayed in cartoon representation with Glc-1P and Ac-CoA in stick and coloured in yellow and cyan, respectively. The *Fo-Fc* omit map contoured at 3σ is displayed around both ligands. (A) Side view, (B) Top view and (C) Bottom view of the E:Ac-CoA:Glc-1P ternary complex.

The image in Figure 65 (A) shows the presence of a metal ion, in this crystal structure the metal is Mg^{2+}, at the threefold axis of symmetry in the trimer, which is in line with previous observations for related protein folds. The top view of the GlmU trimer (Figure 65 (B)) displays the N-terminal regions of each monomer of the trimer, which form the three uridylyltransferase domains. The active sites of these uridylyltransferase domains are formed exclusively from one monomer and are not reliant on trimerisation to form the catalytic pocket. It is clear from this image that the uridylyltransferase domain actives sites are unliganded. Figure 65 (C) highlights that there are three acetyltransferase active sites, which are all simultaneously occupied with Ac-CoA and Glc-1P.

To further focus on the architecture of the GlmU acetyltransferase active site, a close-up is presented in Figure 66 that shows the binding orientations of both Ac-CoA and Glc-1P.

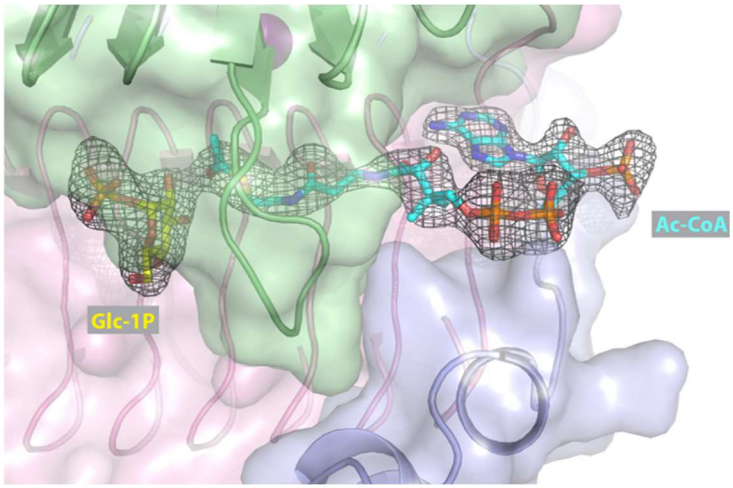


Figure 66. Glc-1P and Ac-CoA in a GlmU acetyltransferase active site.

These experiments were carried out and analysed by Stephane Mouilleron, Structural Biology Platform, The Francis Crick Institute. A close-up image of one of the three GlmU acetyltransferase active sites to demonstrate the orientations of Glc-1P and Ac-CoA ligands in the E:Ac-CoA:Glc-1P ternary complex. The *Fo-Fc* omit map contoured at 3σ is displayed around both ligands.

The image in Figure 66 demonstrates that the binding sites for acyl-CoA and amino sugars are proximal to one another but not overlapping and that a loop, is positioned over the acyl thioester of Ac-CoA. This loop, which has been shown to be disordered in previous substrate-free structures, brings Tyr398 into proximity with Ac-CoA allowing an interaction between its backbone carbonyl and the nearest pantetheine nitrogen (N4P) of Ac-CoA[245]. The image in Figure 66 also shows that all the monomers in the trimer contribute residues to each of the GlmU acetyltransferase active sites.

The binding of Glc-1P to residues in a GlmU acetyltransferase active site if further explored in Figure 67, as well as comparative overlay of the E:Ac-CoA:Glc-1P ternary complex with a previously determined E:CoA-SH:GlcN-1P crystal structure.

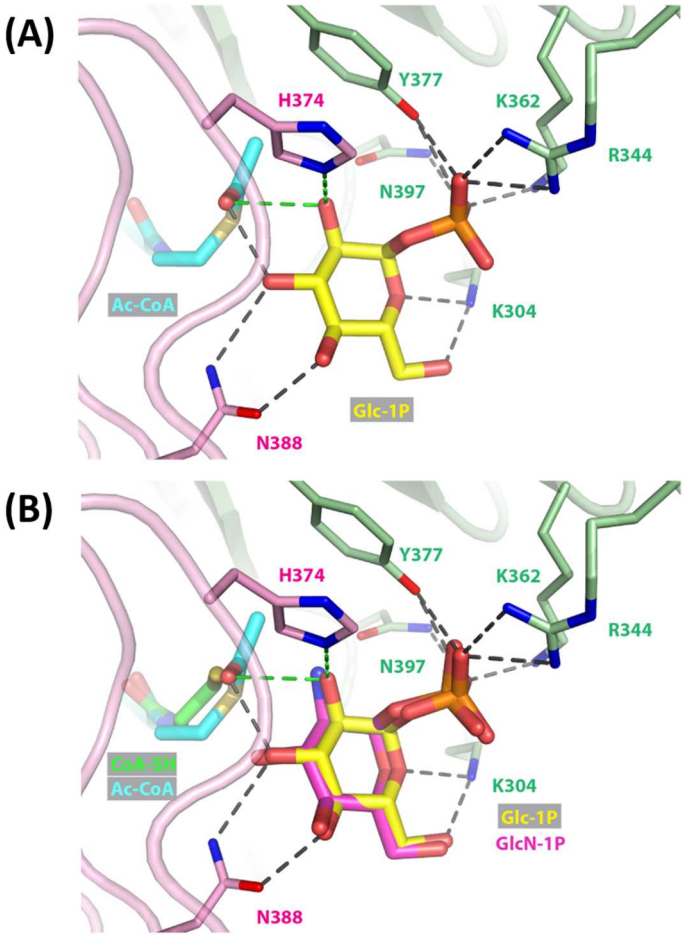


Figure 67. Glc-1P binding site within a GlmU acetyltransferase active site.

These experiments were carried out and analysed by Stephane Mouilleron, Structural Biology Platform, The Francis Crick Institute. The image depicts the Glc-1P/GlcN-1P binding site in an acetyltransferase site of GlmU. For clarity, only the residues contributing to Glc-1P binding are displayed as sticks. The hydrogen bonds formed with Glc-1P atom O2 are displayed in green. (A) Is the same as (B) with GlcN-1P and CoA-SH from the crystal structure 3ST8 superimposed.

The binding mode of Glc-1P to GlmU is very similar to what has previously been reported for GlcN-1P[245]. Glc-1P binds in a pocket that is proximal to the *re* face of the planar acetyl group of Ac-CoA and interacts with three basic residues; Arg344, Lys362 and Lys403, as well as the side chains of Asn397 and Tyr377. In addition, two hydrogen bonds are made with the sidechain of Asn388 belonging to a second monomer. The interactions formed by Lys403 and Asn397 are a result of the ordering of a mobile loop upon Ac-CoA binding, which is highlighted in Figure 68.

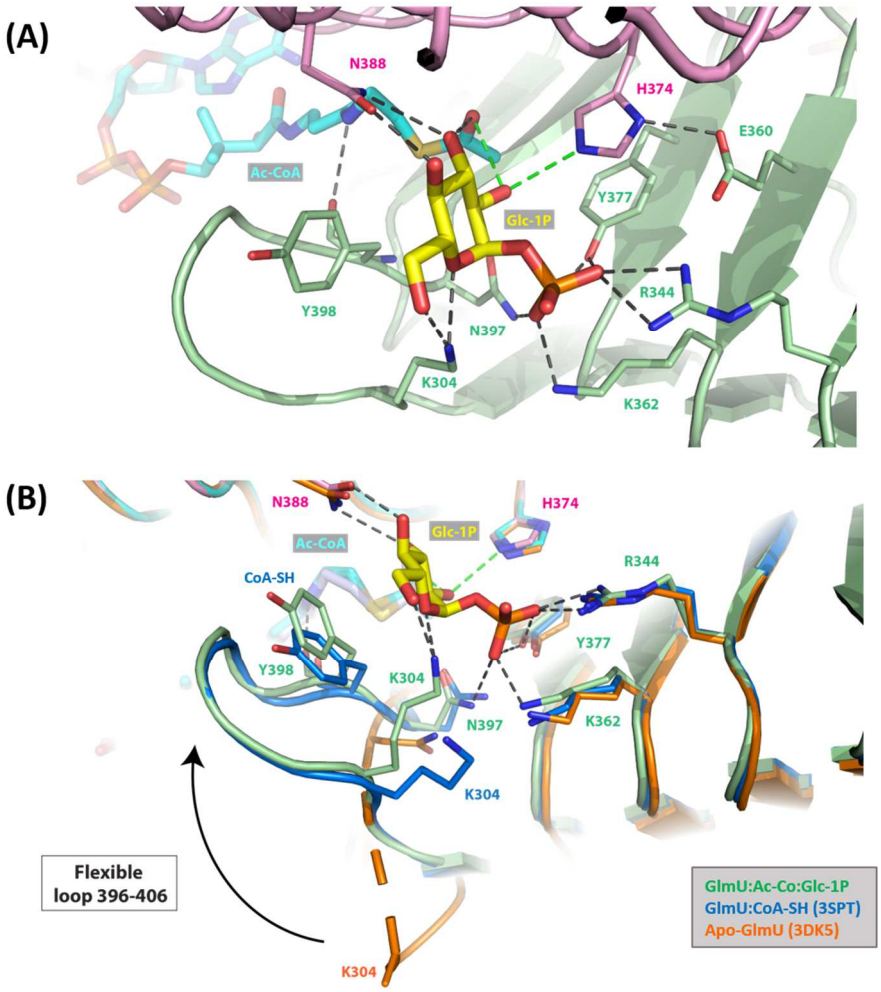


Figure 68. GlmU acetyltransferase active site interactions and conformational changes.

These experiments were carried out and analysed by Stephane Mouilleron, Structural Biology Platform, The Francis Crick Institute. Images showing a wider perspective of the GlmU acetyltransferase active site architecture. (A) Close up view of the catalytic dyad (His374-Glu360), which is proposed to act as the general base in the acetyltransferase reaction. (B) A superposition of two previously determined GlmU structures; the apoenzyme (pdb 3DK5) and a complex with CoA-SH (pdb 3SPT), to illustrate the effect of forming the holoenzyme (with binding of CoA-SH or acyl-CoA) on the ordering of the flexible loop between reside 396 - 406.

An additional interaction enabled by stabilisation of the disordered loop is contributed by the phenolic ring of Tyr398, which upon acyl-CoA binding is brought into proximity with and subsequently stacks with the hexose ring of Glc-1P[424]. Importantly, the Glc-1P 2-hydroxyl group forms two hydrogen bonds, one with the nitrogen Nε2 of the imidazole group of His374 and a second one with the carbonyl group of the Ac-CoA acetyl group. Interestingly, the side chain of the acidic Glu360 forms a hydrogen

bond with His374, which is likely to result in the alignment of the His374 side chain by restricting its rotation, as well as polarisation of the imidazole group by stabilising the positive charge on Nδ1. This interaction is likely to contribute to the local environment modulation of the pK_a of the imidazole allowing it to act as a general base[147].

These X-ray crystallography data also confirm that Mg^{2+} is bound to the GlmU homotrimer at the three-fold crystallographic axis of symmetry. The Mg^{2+} atom in this ternary complex structure (pdb 6GE9) is coordinated by two residues, Ser392 and Asp417 (Figure 69).

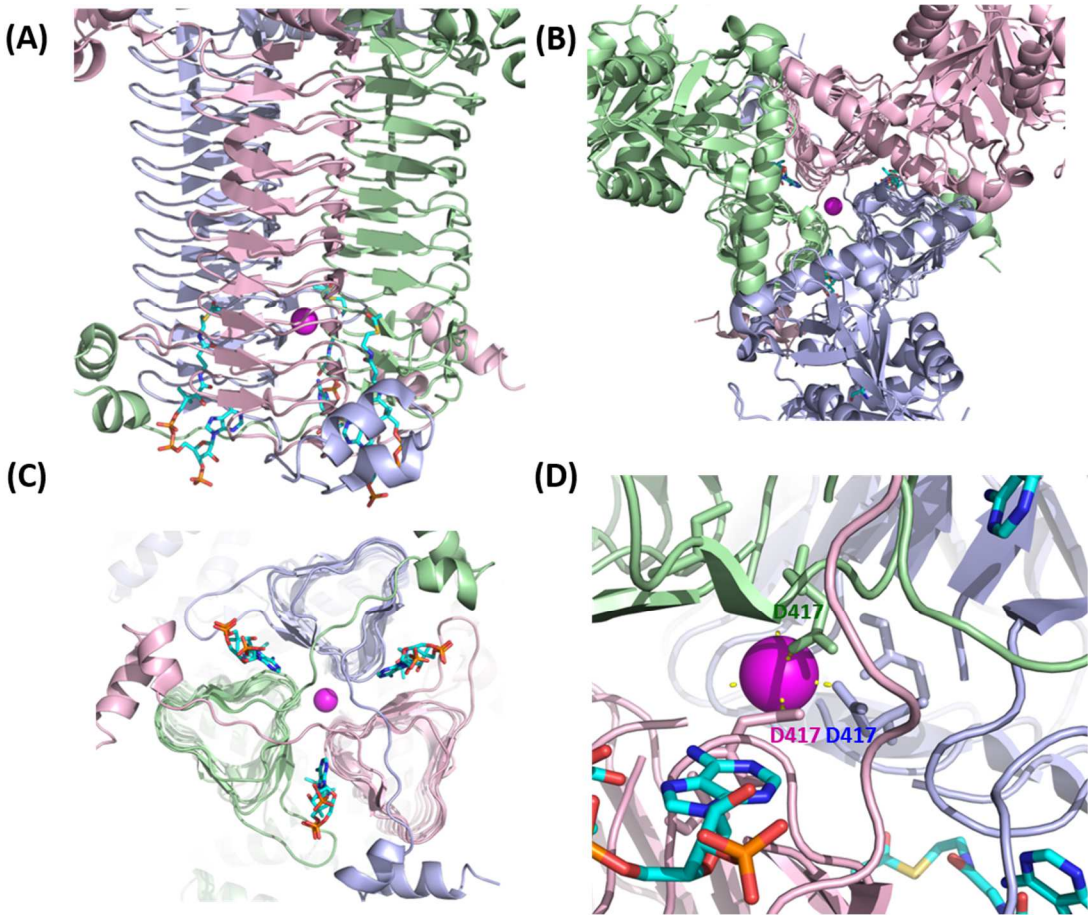


Figure 69. GlmU acetyltransferase domain Mg^{2+} binding.

These experiments were carried out and analysed by Stephane Mouilleron, Structural Biology Platform, The Francis Crick Institute. Images showing a wider perspective of the binding of Mg^{2+} to the GlmU acetyltransferase domain. (A) Side view of the L*β*H timer with the Mg^{2+} bound between three monomers. (B) A bottom view of the acetyltransferase domain trimer with Mg^{2+} bound on the three-fold axis of symmetry. (C) A top view of the acetyltransferase domain trimer with Mg^{2+} bound on the three-fold axis of symmetry. (D) A close-up view of the Mg^{2+} interactions made with the Asp417 residues of each monomer of the L*β*H timer. Ac-CoA is shown in

all of the images to give perspective on the proximity of the Mg^{2+} binding site to the acetyltransferase active site.

The determination of this novel, dead-end analogue ternary complex has enabled a greater understanding of the relative orientations of the residues in the GlmU acetyltransferase active sites, in relation to Ac-CoA and Glc-1P, which, due to the high similarity to the native substrate, GlcN-1P, allows a close approximation of the catalytic form of the complex.

3.13 Proposed kinetic mechanism of GlmU acetyltransferase activity

The results presented in Chapter 3 provide evidence for the proposal of a kinetic mechanism for the GlmU catalysed acetyltransferase activity, which is presented in Scheme 12.

Scheme 12. Proposed kinetic mechanism for *M. tuberculosis* GlmU acetyltransferase activity.

The proposed kinetic mechanism in Scheme 12 is a sequential ordered bi-bi kinetic mechanism, which is supported by the observations from initial velocity pattern experiments, a combination of competitive and uncompetitive inhibitor patterns and finally, the structural biology data for X-ray crystallography studies that captured a dead-end inhibited ternary complex.

Further investigations were subsequently carried out, documented in Chapter 4 , to better understand and characterise the chemical mechanism of the GlmU acetyltransferase activity.

Chapter 4. Characterisation of the chemical mechanism of GlmU catalysed acetyl transfer

The data presented in the previous chapter provided the evidence to propose a kinetic mechanism, a sequential ordered bi-bi kinetic mechanism (Scheme 12), that describes the order of substrate binding and product release. Subsequently, to better understand *M. tuberculosis* GlmU acetyltransferase activity, a series of studies, the results of which are documented in this chapter, were carried out to interrogate the chemical mechanism. The aim of these studies is to; primarily, identify the resides in the acetyltransferase active site that are responsible for substrate recognition and catalysis, and subsequently, elucidate complete catalytic mechanism of GlmU-catalysed acetyl transfer. Most of the studies carried out to understand the *M. tuberculosis* GlmU acetyltransferase chemical mechanism are predominantly based on structural biology approaches, however, the very nature of these precludes any understanding of the temporal aspects of catalysis. Initially, a detailed examination of the pH dependence of the acetyltransferase Michaelis-Menten parameters of GlmU, were carried out to identify the acid-base residues involved in both substrate recognition and catalysis. Additionally, to understand and complement the pH-rate profile experiments, 1D NMR studies were carried out to measure the pK_a ionisable groups on GlcN-1P. The findings of the NMR studies will contribute to assignment of key pK_a values, revealed by the acetyltransferase pH-rate profile experiments. Prior to the initiation of this project, His374 had been suggested as being essential for catalysis, indeed previous mutation of this residue lead to loss of acetyl transfer under the experimental conditions[245]. To further study the role of His374 in catalysis, site-directed mutagenesis was carried out to change His374 to an Ala and, subsequently, pH-rate profile experiments were undertaken to provide a comparative data set, which will contribute to assignment of the pK_a values identified in the WT pH-rate profile. Additionally, key amino acid residues for GlcN-1P binding, identified from the structural biology studies in chapter 3; Arg344, Lys362 and Tyr377, were systematically changed by site-directed mutagenesis to Ala and the activity of the resulting mutant proteins were characterised. Additional pH-rate studies were carried out with each of the mutant enzymes. Finally, solvent kinetic isotope effect studies, which involve the isotopic substitution of the solvent water

hydrogens to deuterium, were undertaken, utilising WT GlmU, to identify if solvent-derived proton transfer(s) is or are involved in acetyl transfer. All equations used in the analysis of data in this chapter can be found in the Materials and Methods chapter, section 2.2.30.

4.1 Acetyltransferase pH-rate profile

The effect of varying pH on GlmU acetyltransferase Michealis-Menten kinetic parameters, was studied to assess the role of general acid-base chemistry in substrate binding and catalysis. To determine the pH-rate profile of GlmU-catalysed acetyl transfer, a range from pH 6.0 to 10.0, was utilised for these studies. There are several issues involved in configuring pH-rate experiments that need to be addressed prior to initiation of the actual study. The experimental challenges involved in varying the pH of an assay can include; selection of a suitable buffer, or mixture of buffers that maintain ionic strength in the assay, stability of the enzyme tertiary structure and maintenance of the sensitivity of the chosen assay detection methodology. To address these challenges a series of control experiments were systematically carried out, prior to investigating the effect of pH on the rate of acetyl transfer.

Selection of a suitable buffer that could be used to maintain assay pH conditions across the desired range, while maintaining ionic strength, was critical and led to identification of a “tri-buffer” system that had previously been described by Ellis and Morrison[398] (Methods section 2.2.20). An initial experiment was carried out to assess the applicability of the tri-buffer system for kinetically monitoring acetyltransferase activity using the DTP absorbance assay (Figure 70).

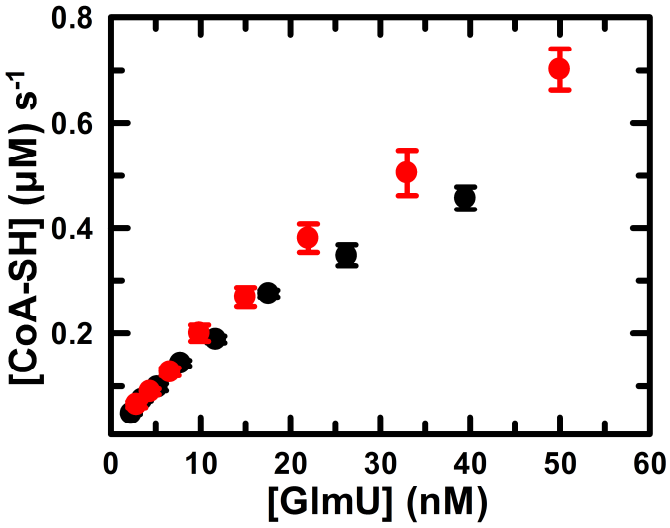


Figure 70. Comparison of acetyltransferase activity in two different assay buffers.

Enzyme titrations in either the HEPES-based buffer used for all early absorbance assays (●) or the tri-buffer system (●) prior to initiation of acetyltransferase activity by the addition of 200 μM Ac-CoA, GlcN-1P and 400 μM DTP. The buffers were both set to pH 7.5. For each enzyme concentration, the data are a mean of either 2 or 4 discrete replicates.

The initial validation of the tri-buffer, adjusted to pH 7.5, for use with the DTP absorbance assay was comparable with the HEPES-based buffer (Methods section 2.2.8) that has been used for the majority of kinetic studies (Figure 70). To assess the sensitivity of the DTP absorbance assay to CoA-SH, a standard curve of the acetyltransferase product was prepared in tri-buffer adjusted to pH values between 6 and 10. (Figure 71).

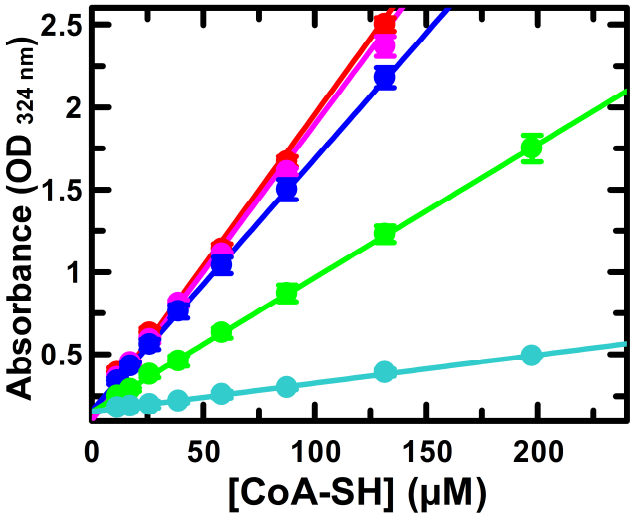


Figure 71. CoA-SH standard curves in tri-buffer adjusted to varying pH.

CoA-SH titrations in tri-buffer, adjusted to the following pH values; 6 (●), 7 (●), 8 (●), 9 (●), and 10 (●). Each of the points in the plot are a mean of three replicates. The data for each pH condition were fitted to a linear regression model.

The sensitivity of DTP to varying concentrations of CoA-SH does not differ significantly between pH 6.0 to 8.0 but reduces as the buffer becomes more basic. This change in sensitivity of DTP at basic pH conditions is well documented[421] under several assay conditions, and is primarily due to a sharp decrease in the molar absorption coefficient of 4-TP. To address this, and to control for experimental variables, product standard curves were carried out on every test occasion to determine the concentration of CoA-SH from the increase in absorbance.

GlmU protein stability across the range of desired pH conditions was assessed by monitoring the change in the enzyme T_m by DSF. GlmU was incubated for 10 minutes in tri-buffer, adjusted to pH spanning the desired range of 6.0 to 10.0, prior to addition of 10x SYPRO Orange and subsequent thermal denaturation (Figure 72, Table 18).

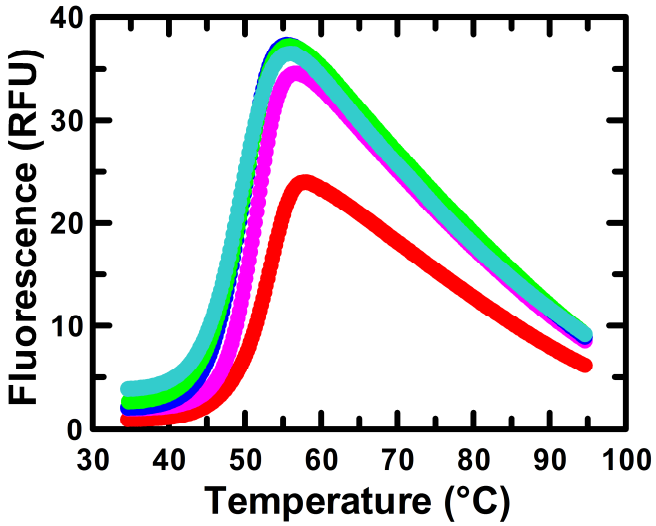


Figure 72. The effect of pH on GlmU T_m.

DSF thermal melting profiles of 2 µM GlmU in 10x SYPRO Orange, prepared in tri-buffer, adjusted to the following pH values; 6 (●), 7 (●), 8 (●), 9 (●), and 10 (●). Each of the points in the plot are a mean of six replicates. The data for each pH condition were fitted to Eqn. (19) and the Tm data for each condition are presented in Table 18.

Table 18. GlmU Thermal Stability in varying pH conditions [a].

pH	T_m (°C)
6	51.23 ± 0.37
7	51.40 ± 0.13
8	48.81 ± 0.05
9	49.06 ± 0.08
10	48.64 ± 0.02

[a] Thermal stability experiments performed utilizing the same buffer adjusted to the desired pH using either 6 M HCl or 10 M NaOH.

The data presented in Figure 72 demonstrates that under all pH conditions tested GlmU is likely to have an intact tertiary structure, due to the typical thermal unfolding profiles. The data in Table 18 demonstrates that GlmU T_m varied by 2.5 °C between the pH conditions assessed, which was indicative that the protein is affected by the buffer conditions. The GlmU T_m is most stable at pH 6.0 and least stable at pH 10,

however, due to the likely intact tertiary structure (Figure 72) further control studies were carried out. The thermal profile of GlmU in pH 6.0 conditions had a 20 % lower fluorescence maximum, which could be indicative of a reduction of total protein. To further characterise the catalytic activity of the GlmU acetyltransferase domain, enzyme titrations were carried out in buffer across the pH range of 6.0 to 10.0, prior to initiation of catalysis by the addition of 200 µM Ac-CoA, GlcN-1P and 400 µM DTP (Figure 73).

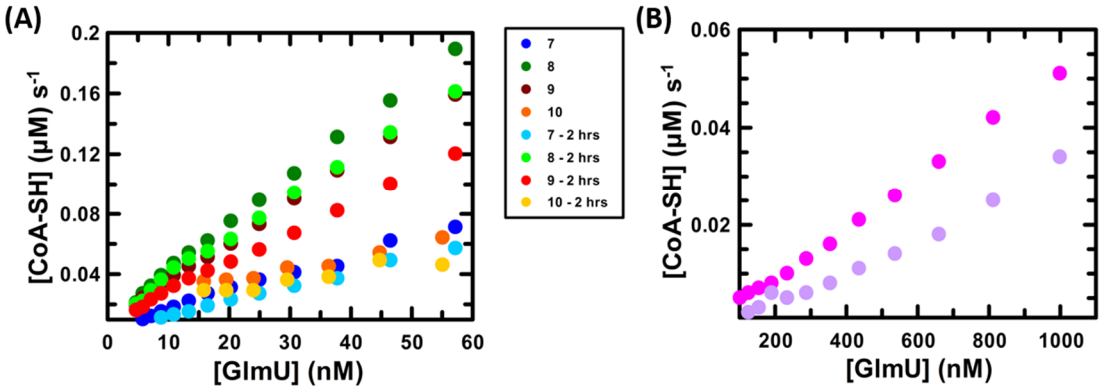


Figure 73. GlmU acetyltransferase activity in different pH conditions.

Enzyme titration in tri-buffer adjusted to different pH conditions, acetyltransferase activity initiated by addition 200 µM Ac-CoA, GlcN-1P and 400 µM DTP and CoA-SH formation monitored using the change in absorbance at 324 nm. (A) Comparison of the acetyltransferase activity of GlmU titrations in buffer adjust to pH values from 7 to 10, after either a 5 mins or 2 hrs (at 4 °C) incubation. (B) Comparison of acetyltransferase activity of GlmU titrations in buffer adjust to pH 6, after either a 5 mins (●) or 2 hrs (●) incubation. Each of the points in the plot are a mean of three replicates.

The data in Figure 73 demonstrates that it was possible to monitor linear GlmU acetyltransferase activity in tri-buffer adjusted to pH conditions between 6.0 to 10.0. It is apparent that there was a modest loss of GlmU acetyltransferase activity between the 5 mins and 2 hrs incubations in the different pH conditions, which was at most 25%. GlmU acetyltransferase activity between the pH conditions is different when the response of the same enzyme concentrations was compared. A 32-fold difference between the acetyltransferase activity of 35 nM GlmU was observed between the pH 6 and pH 8 buffer conditions. These data could be indicative of protein instability, which was observed when monitoring GlmU T_m in Table 18. However, this difference in activity could also be due to the change in ionisation state

of either the enzyme or reactants in different pH conditions. To better understand the effect of varying pH on GlmU acetyltransferase activity, Michaelis-Menten parameters for both Ac-CoA and GlcN-1P were determined, using the DTP absorbance assay, at 0.25 increments pH between pH 6 and 10 (Figure 74).

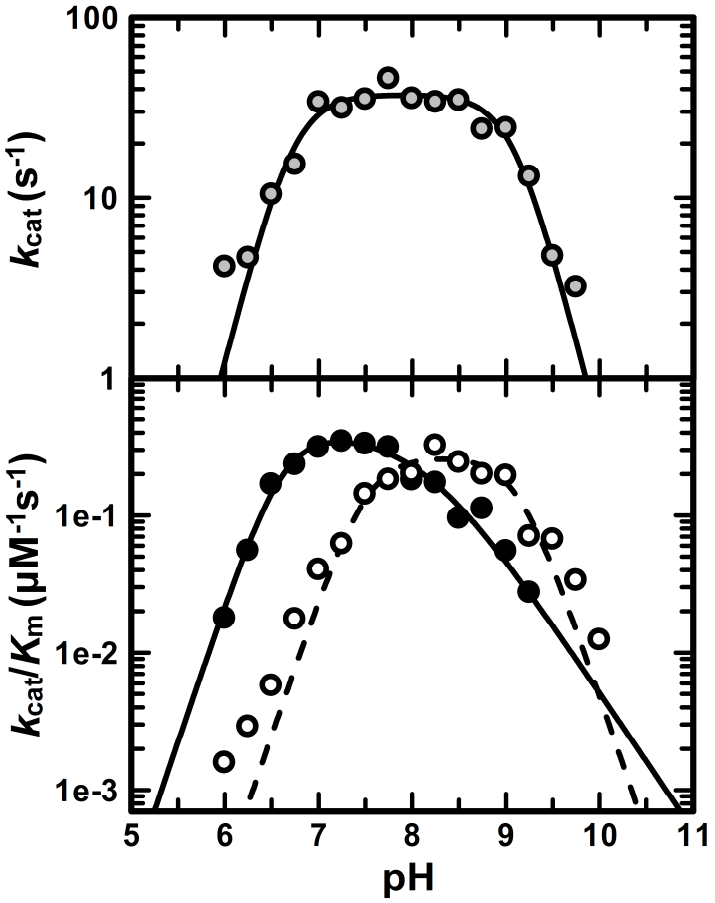


Figure 74. pH dependence of GlmU acetyltransferase kinetic parameters.

The k_{cat} and K_m values were determined at each pH by varying concentration of one substrate at a fixed, saturating concentration of the second substrate. The k_{cat} data is a mean of the Ac-CoA and GlcN-1P values and are represented by grey filled circles and the best fit of the data, to Eqn. (11), is represented by ─. Ac-CoA experimental data are represented by (●) and the best fit of the data, to Eqn. (12), is represented by ─. The GlcN-1P experimental data are represented by (○) and the best fit of the data, to Eqn. (11), is represented by - - -.

The data presented in Figure 74 shows a bell-shaped pH profile for demonstrated that there was a pH dependence of k_{cat}, which, by the shape of the k_{cat} plot, revealed both acid- and base-assisted catalysis. The bell-shaped curve of the k_{cat} plot versus pH, had slopes of +2 and -2 in the acidic and basic limbs, respectively, which was best fitted to a model describing the presence of two acidic non-resolvable and two

basic non-resolvable ionisable groups (Eqn. (11)). The data when fitted to Eqn. (11) yielded pK_a values of 6.7 ± 0.1 and 9.0 ± 0.1, respectively for the groups that need to be deprotonated (general base) and protonated (general acid) for maximal activity. Similarly, the plot of pH versus $k_{cat}/K_{m,\ GlcN\text{-}1P}$ was bell-shaped, with slopes of +2 and -2 in the acidic and basic regions, respectively. These data were best fitted to Eqn. (11), describing the role of two acidic nonresolvable and two basic non-resolvable ionisable groups, and yielded pK_a values of 7.5 ± 0.1 and 9.1 ± 0.1, respectively for the acidic and basic groups. Finally, the bell-shaped plot of pH versus log $k_{cat}/K_{m,\ Ac\text{-}CoA}$ had slopes of +2, at low pH and -1, at high pH, which was best fit to an equation describing the role of two nonresolvable acidic groups and one basic group (Eqn. (12)). The $k_{cat}/K_{m,\ Ac\text{-}CoA}$ data when fitted to Eqn. (12) yielded pK_a values of 6.6 ± 0.1 and 8.1 ± 0.1, respectively, for the two non-resolvable acidic groups and one basic group.

4.2 GlcN-1P NH_3 pK_a determination using 1D NMR

To identify the ionisable groups involved in GlmU acetyltransferase activity, 1D NMR studies were carried out to determine the pK_a value of the GlcN-1P amine group, which may represent one of the ionisable groups observed in the pH-rate studies (k_{cat} or $k_{cat}/K_{m,\ GlcN\text{-}1P}$). To determine the ionisation state of GlcN-1P in solution, NMR chemical shifts of two ^{1}H atoms, were monitored over a range of pH from 5.5 to 11.5, using the methodology described in Methods section 2.2.23 (Figure 75).

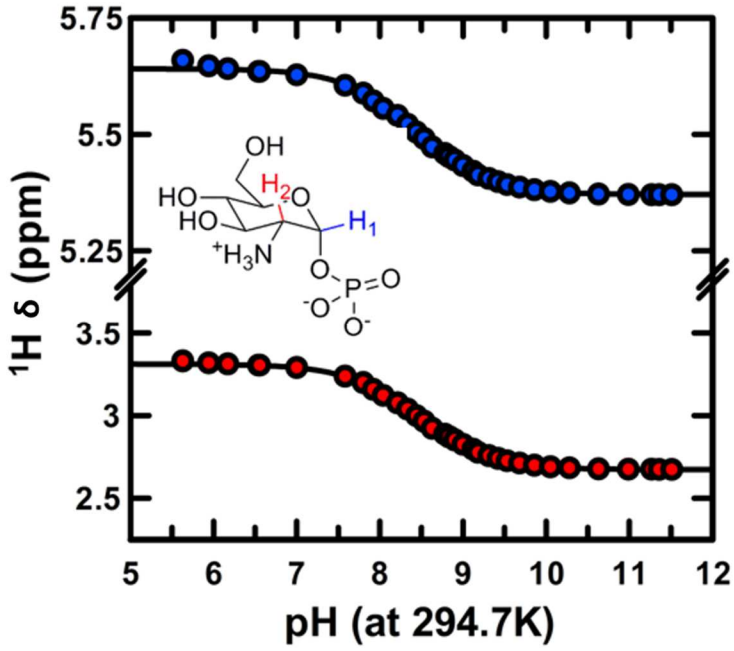


Figure 75. Determination of the ionisation states of GlcN-1P in solution using 1D NMR.

These experiments were carried out and analysed by Cesira de Chiara, Mycobacterial Metabolism and Antibiotic Research Laboratory, The Francis Crick Institute. Acidic and basic preparations of GlcN-1P were titrated to determine a pK_a for the protonated amine group. The ^{1}H peak intensities were monitored of CH(2)-NH_{3+} (experimental data represented by, ●) and CH(1)-PO_4H_2 (experimental data represented by, ●) were monitored. The chemical shift for CH(2)-NH_3^+ was 3.33 ppm, while the CH(1)-PO_4H_2 shift was 5.65 ppm. The data was fit to Eqn. (21) and is represented by ─. The structure of GlcN-1P is inset and the CH(2)-NH_3^+ and CH(1)-PO_4H_2 highlighted in red and blue, respectively.

The results presented in Figure 75, depicts ^{1}H chemical shift data fitted to Eqn. 19, which allowed determination of value of 3.33 ppm for CH(2)-NH_3^+ and a value of 5.65 ppm for CH(1)-PO_4H_2. The data was fitted to Eqn. (21), which yielded a pK_a for CH(2)-NH_3^+ of 8.43 ± 0.01 and a pK_a for CH(1)-PO_4H_2 of 8.44 ± 0.01.

4.3 Preparation and characterisation of GlmU H374A

To further understand and assign the ionisable groups identified in the GlmU acetyltransferase pH-rate profiles, SDM was carried out to change His374 to an Ala residue. The position of the His374 in the acetyltransferase active site appears to be optimal for this residue to act as the catalytic residue in the (Figure 67). Furthermore, mutagenesis of this residue to Ala has previously been shown to

reduce *M. tuberculosis* GlmU-catalysed acetyl transfer by 90% in comparison to the WT, however, mutation of Asp397 to Ala also reduced acetyltransferase by 95%[245]. The mutation of His374 to Ala was carried out using the Phusion SDM kit, as per the method described in Methods section 2.2.5. The GlmU H374A SDM was confirmed as successful by Sanger sequencing of the mutated plasmid, prior to protein expression and purification[425]. The mutant GlmU, referred to as GlmU H374A from hereon in, was prepared as per the protocol described in Methods section 2.2.6. The purified GlmU H374A was analysed by reducing SDS-PAGE gel electrophoresis (Figure 76) and protein concentration was determined by Bradford protein assay, which was used to calculate the total protein yield.

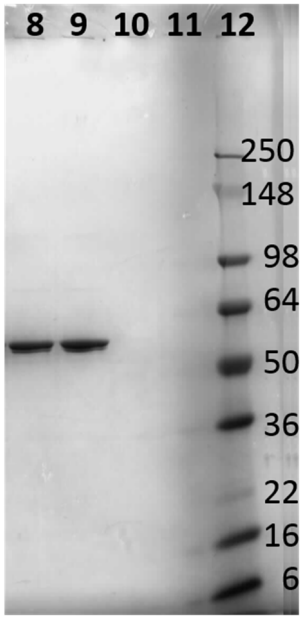


Figure 76. SDS-PAGE gel electrophoresis analysis of His_6-TEV-GlmU H374A.
SDS-PAGE analysis of the final purification product of His_6-TEV-GlmU H374A. The gel layout is as follows; lane 8 & 9 post-SEC, concentrated GlmU H374A (2 µg protein loading mass) and lane 12 is the gel MWt marker.

The protein concentration was determined as 4.98 mg/ml, which meant that a total of 49.8 mgs of protein were purified from 18 grams of cell pellet. The protein accurate mass was determined by ESI LC-MS as 53.4 kDa, which is consistent with N-terminal methionine deletion. These data, in tandem with the positive DNA sequencing data, confirm the homogeneity and mass of the protein purified from the *E. coli* cell pellet as GlmU H374A.

DSF thermal stability studies were utilised to determine that the GlmU H374A tertiary structure is intact and to ascertain whether there was any change in T_m when compared to WT GlmU (Figure 77).

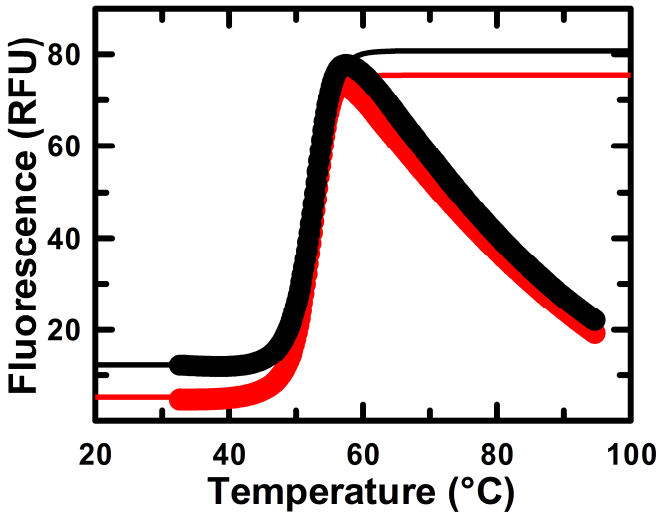


Figure 77. The effect of the H374A mutation on GlmU T_m.

DSF thermal melting profiles of either 2 µM GlmU (●) or GlmU H374A (●) in 10 x SYPRO Orange. Each of the points in the plot are a mean of six replicates. The mean data were fitted to Eqn. (19) and the T_m data for each condition are presented below.

The data presented in Figure 77 when fitted to Eqn. (19), yielded a WT T_m = 52.15 ± 0.03 °C, and a H374A T_m = 52.04 ± 0.04 °C, which demonstrates that the mutation has no effect on the thermal stability of GlmU. Further DSF studies were carried out to ascertain whether there was any modulation of GlmU H374A T_m in the presence of titrations of either CoA-SH or GlcN-1P (Figure 78).

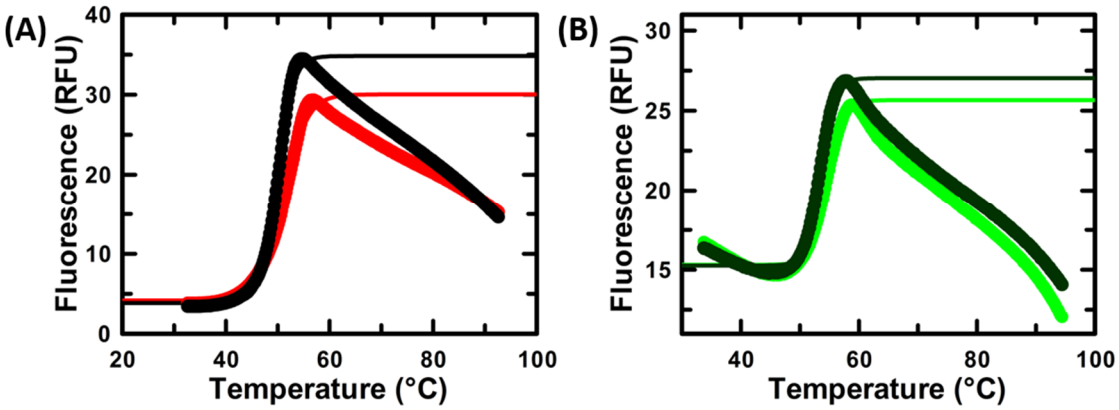


Figure 78. The effect of CoA-SH and GlcN-1P on GlmU H374A T_m.

DSF thermal melting profiles of GlmU H374A, using 10 x SYPRO Orange, in the presence of a concentration titration of either CoA-SH or GlcN-1P. (A) 2 μM GlmU H374A incubated with a titration of CoA-SH, exemplified on the plot by either 65 μM (●) or 1000 μM (●). (B) 2 μM GlmU H374A in the presence of 1000 μM CoA-SH, prior to an incubation with a titration of GlcN-1P, exemplified by either 95 μM (●) or 3000 μM (●) GlcN-1P. Each of the points in both plots are a mean of six replicates. The mean data were fitted to Eqn. (19) and the T_m data for each condition are presented below.

The data in Figure 78 demonstrated that GlmU H374A T_m was stabilised by 1.7 °C when incubated with 1000 μM CoA-SH, which was the same stabilisation that was observed the wild type protein (data now shown). Additional studies were carried out using a co-incubation of GlmU H374A and 1000 μM CoA-SH, prior to adding a titration of GlcN-1P, which led to an additional stabilisation of GlmU H374A T_m by 1.2 °C. The additional stabilisation of GlmU H374A T_m in the presence of GlcN-1P was previously observed with the WT protein, co-incubated with high concentrations of either Ac-CoA or CoA-SH, in the presence of a concentration of Glc-1P, which stabilised the GlmU T_m by a maximum of 2.2 °C.

GlmU H374A acetyltransferase activity was initially assessed by an enzyme titration in the presence of 200 μM Ac-CoA and GlcN-1P, in assay buffer adjusted to pH 7.5, using 400 μM DTP to monitor the formation of CoA-SH by absorbance change at 324 nm. The WT enzyme was titrated, and acetyltransferase activity was monitored to act as a control for the SDM protein (Figure 79).

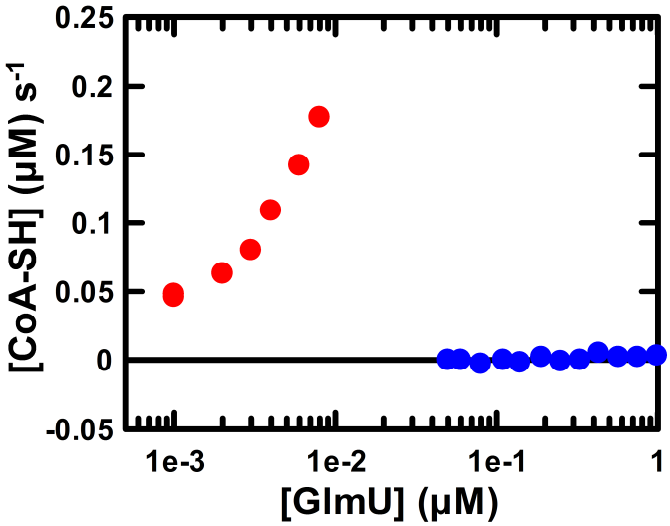


Figure 79. GlmU titration to monitor acetyltransferase activity to compare WT and H374A enzyme preparations.

GlmU enzyme titration to compare the acetyltransferase activity of WT (●) and H374A (●) preparations of protein. GlmU was titrated in assay buffer adjusted to pH 7.5, prior to initiation of the acetyltransferase activity by 200 µM Ac-CoA and GlcN-1P before monitoring CoA-SH formation using 400 µM DTP and measuring change in absorbance at 324 nm.

The mutation of His374 to Ala in the GlmU acetyltransferase active site, ablates acetyl transfer activity when compared to the WT protein in assay buffer adjusted to pH 7.5. The data presented in Figure 79 demonstrated that none of the GlmU H374A concentrations tested had any significant acetyltransferase activity, however, the control GlmU WT concentrations used produced CoA-SH at a similar rate to previously determined levels (Figure 27). Further studies will be carried out to understand whether varying the concentration of the acetyltransferase substrates or pH conditions of the assay buffer have any effect on the acetyl transfer activity of GlmU H374A.

4.4 GlmU H374A pH-rate profile

To confirm the role of His374 in the GlmU-catalysed acetyl transfer reaction, pH-rate profile studies were undertaken by monitoring acetyltransferase activity at various Ac-CoA and GlcN-1P concentrations at different pH conditions (Figure 80).

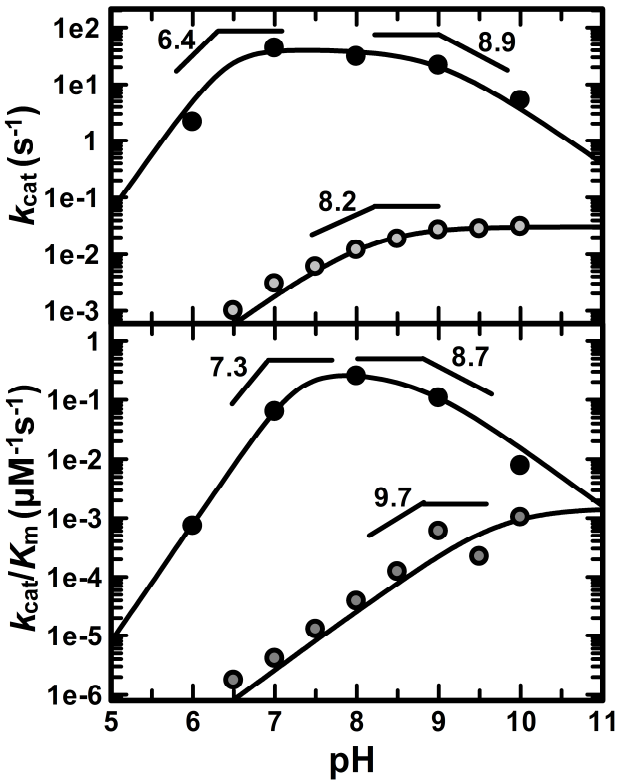


Figure 80. pH-rate profiles of GlmU and GlmU H374A acetyltransferase GlcN-1P kinetic parameters.

The GlmU acetyltransferase k_{cat} and K_m values were determined at each pH by varying concentration of GlcN-1P at fixed, saturating concentrations of Ac-CoA. The GlcN-1P data relating to WT GlmU k_{cat} and k_{cat}/K_m data is represented by black filled circles and the best fit of the data, to Eqn. (12), is represented by ─. The GlcN-1P data relating to GlmU H374A k_{cat} and k_{cat}/K_m data is represented by grey filled circles and the best fit of the data, to Eqn. (9), is represented by ─.

GlmU H374A acetyltransferase activity is significantly slower compared to enzyme than WT GlmU, displaying a pH independent plateau which is approximately three orders of magnitude lower than that of the WT GlmU (Figure 80). The pH dependence of GlmU H374A k_{cat} reveals a single ionisable group that must be deprotonated for maximal activity (slope of +1). These data were best fitted to Eqn. (9) that describes the involvement of one basic ionisable group (pK_a = 8.2 ± 0.1). Similarly, the plot of pH versus GlmU H374A $k_{cat}/K_{m,\ GlcN\text{-}1P}$ showed a single ionisable group (slope of +1), which was best fitted to Eqn. (9) that allowed determination of a value of 9.7 ± 0.5 for the basic group. The acetyltransferase pH-rate profile data for GlmU WT was generated in parallel with the GlmU H374A study. The WT control study, utilised GlcN-1P as the varied substrate in the presence of saturating Ac-CoA,

was best fitted to the equation for two nonresolvable acidic groups and one basic group (Eqn. (12)). The GlmU WT pH-rate profile data when fitted to Eqn. (12) yielded pK_a values for $k_{cat,\ GlcN\text{-}1P}$ of 6.4 ± 0.3 and 8.9 ± 0.3, and pK_a values for $k_{cat}/K_{m,\ GlcN\text{-}1P}$ of 7.3 ± 0.02 and 8.7 ± 0.04, in both cases, respectively for the acidic and basic groups.

In addition to determining the GlmU H374A catalysed GlcN-1P Michaelis-Menten parameters in varying pH conditions, the Ac-CoA steady state parameters were determined, utilising the same pH range, along with WT controls (Figure 81).

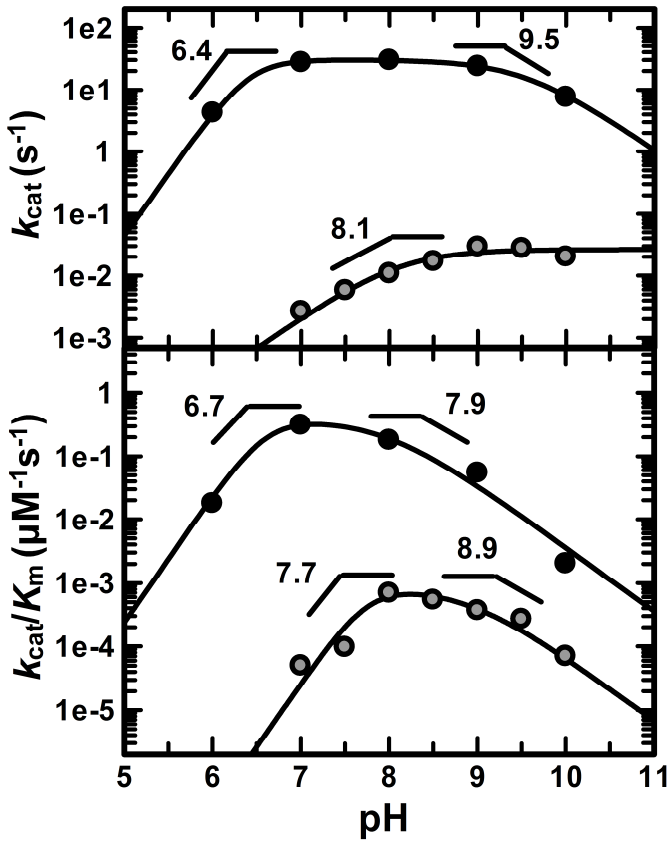


Figure 81. pH-rate profiles of GlmU and GlmU H374A acetyltransferase Ac-CoA kinetic parameters.

The GlmU acetyltransferase k_{cat} and K_m values were determined at each pH by varying concentration of Ac-CoA at fixed, saturating concentrations of GlcN-1P. The Ac-CoA data relating to WT GlmU k_{cat} and k_{cat}/K_m data is represented by black filled circles and the best fit of the data, to Eqn. (12), is represented by ─. The Ac-CoA data relating to GlmU H374A k_{cat} and k_{cat}/K_m data is represented by grey filled circles and the best fit of the data, to either Eqn. (9) or Eqn. (12), respectively, is represented by ─.

As previously observed with the GlcN-1P steady state parameter pH dependence study, GlmU H374A is a significantly slower acetyltransferase enzyme than WT GlmU, displaying a pH independent plateau which is approximately three orders of magnitude lower than WT GlmU (Figure 81). The pH dependence of GlmU H374A k_{cat} reveals a single ionisable group that must be deprotonated for maximal activity (slope of +1). These data were best fitted to Eqn. (9) that describes the involvement of one basic ionisable group, which yielded a pK_a value of 8.1 ± 0.02. The plot of pH versus GlmU H374A log $k_{cat}/K_{m,\ Ac\text{-}CoA}$ showed two nonresolvable acidic ionisable and one basic ionisable group, which was best fitted to Eqn. (12) that allowed determination of pK_a values of 7.7 ± 0.2 and 8.9 ± 0.3. Further acetyltransferase pH-rate profile data for WT GlmU was generated in parallel with the GlmU H374A study. This control study, utilised Ac-CoA as the varied substrate in the presence of saturating GlcN-1P, was best fitted to the equation for two nonresolvable acidic groups and one basic group (Eqn. (12)). The GlmU WT pH-rate profile data when fitted to Eqn. (12) yielded pK_a values for $k_{cat,\ Ac\text{-}CoA}$ of 6.4 ± 0.04 and 9.5 ± 0.04, and pK_a values for $k_{cat}/K_{m,\ Ac\text{-}CoA}$ of 7.7 ± 0.2 and 8.9 ± 0.3, in both cases, respectively for the acidic and basic groups.

The data generated in these studies confirm that the His374 residue is essential for optimal GlmU catalysed acetyl transfer, with k_{cat} reduced by approximately 1000-fold when this residue is replaced by Ala. The pH-rate profiles for k_{cat} and k_{cat}/K_m, with relation to both GlcN-1P and Ac-CoA observed using the GlmU H374A mutant protein are significantly altered from the WT enzyme and provide valuable information to elucidate the role of His374 in GlmU acetyltransferase activity.

4.5 Further GlmU SDM preparation and characterisation

The GlmU acetyltransferase active site residues Arg344, Lys362, Tyr377 and Asn388 were further selected for further investigation, based on the structural information presented in Figure 67, where each of the amino acids were identified to be involved in interactions with GlcN-1P. Additionally, the pK_a values identified in both the k_{cat} and k_{cat}/K_m GlmU acetyltransferase pH rate profiles could potentially be attributed to some of these residues.

The mutations of Arg344, Lys362, Tyr377 and Asn388 to Ala were carried out using the Phusion SDM kit, as per the method described in Methods section 2.2.5. The

SDM changes to the GlmU WT DNA were confirmed by use of Sanger sequencing, which identified that the N388A mutation was unsuccessful[425]. The mutant GlmU proteins, subsequently referred to as either GlmU R344A, GlmU K362A and GlmU Y377A (SDM confirmed by Sanger sequencing), were prepared with a novel high throughput (HT) protocol, described in Methods section 2.2.7. As this purification methodology was unprecedented for GlmU, the WT enzyme, as well as GlmU H374A were prepared in parallel to act as controls in the initial characterisation experiments. The purified GlmU SDM proteins, in addition to the two controls, were analysed using both reducing SDS-PAGE gel electrophoresis (Figure 82), as well as the Caliper microfluidic capillary electrophoresis system, which allowed determination of purity and concentration in the same experiment (Table 19).

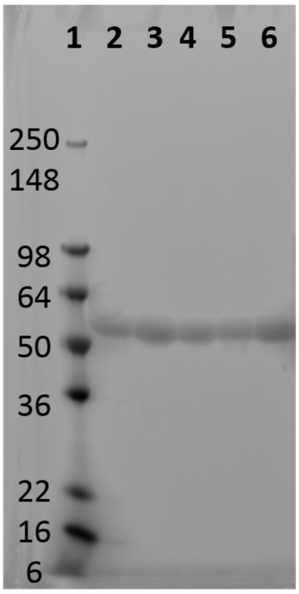


Figure 82. SDS-PAGE gel electrophoresis analysis of His_6-TEV-GlmU SDM mutants.

SDS-PAGE analysis of the final HT purification products of His_6-TEV-GlmU SDM mutants and the WT control. The gel layout is as follows; lane 1 - gel MWt marker, lane 2 - GlmU WT, lane 3 - GlmU R344A, lane 4 - GlmU K362A, lane 5 - GlmU Y377A and lane 6 - GlmU H374A (all 2 µg protein loading mass).

Table 19. Caliper microfluidic capillary electrophoresis analysis of His_6-TEV-GlmU SDM mutants.

GlmU protein	Concentration (mg/ml)	Purity (%)
WT	3.36	91.1
R344A	4.39	94.3
K362A	3.55	91.6
Y377A	2.77	88.1
H374A	3.47	97.1

Protein preparations in a storage buffer comprised of 25 mM HEPES, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$.

In addition, the protein accurate masses were determined by ESI LC-MS and all were measured as 53.4 kDa, which are like previous observations and are consistent with a N-terminal methionine deletion. These data confirm the purity and mass of the proteins purified from the *E. coli* cell pellets as GlmU, or single residue mutant variations of the WT. Importantly, the analyses of these HT protein preparations are consistent, in terms of mass and purity, with previous evaluations of GlmU batches prepared using Äkta column chromatography methods.

To determine whether the preparations of HT GlmU SDM mutants and corresponding controls had intact tertiary structure, DSF thermal stability studies were carried out in the presence of SYPRO Orange (Table 20).

Table 20. Melting temperature determination of different GlmU preparations [a].

GlmU protein	T_m (°C)
WT	51.10 ± 0.04
R344A	52.50 ± 0.05
K362A	50.50 ± 0.04
Y377A	50.51 ± 0.03
H374A	45.15 ± 0.03
WT, previous preparation	51.52 ± 0.04
WT pre-incubated with 1 mM CoA-SH, previous preparation	53.31 ± 0.06
H374A, previous preparation	51.26 ± 0.02

[a] Thermal stability experiments performed utilising an assay buffer comprised of 25 mM HEPES, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$. Each experiment was a mean of 24 individual measurements and utilised 1 µM protein per test. Data were fit to Eqn. (19).

The data presented in Table 20 demonstrates that GlmU T_m did vary between the protein preparations assessed, however, the thermal melting profiles of all the variants tested were indicative of a protein with an intact tertiary structure. The T_m of the GlmU WT preparations were broadly in agreement, which further validates the HT preparation of GlmU. The mutant form of GlmU that varies significantly is the HT H374A variant, which had a T_m 5 °C lower than the previous preparation (Figure 77). These data suggest that the HT GlmU H374A preparation was partially unfolded or destabilised by the change in the purification process (difference between conventional chromatography and HT methodology) and will not be utilised for any further studies. Conversely, the GlmU R344A T_m was stabilised by between 1 to 1.4 °C, when compared to either of the two WT preparations, which indicates that the residue change to Ala has a stabilising effect on the GlmU tertiary structure.

The effect of the three GlmU acetyltransferase active site mutations, in addition to the change purification methodology, were further assessed by monitoring the

catalytic activity. To carry out this study, enzyme titrations of each of the preparations, listed in Table 20, were carried out in assay buffer adjusted to pH 7.5, prior to initiation of catalysis by the addition of 200 µM Ac-CoA, GlcN-1P and 400 µM DTP and evaluation of acetyl transfer by change in absorbance at 324 nm (Figure 83).

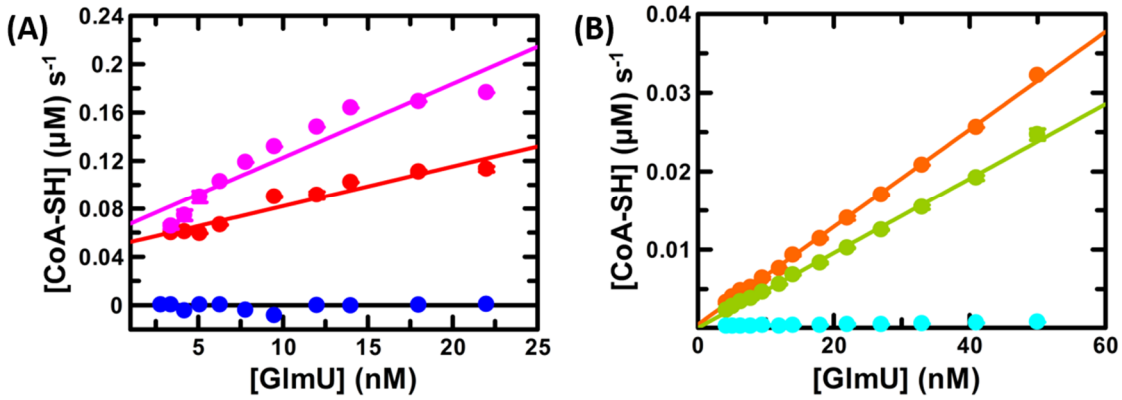


Figure 83. Enzyme titration to compare acetyltransferase activity of SDM mutants and different preparations of GlmU.

GlmU enzyme titrations in assay buffer adjusted to pH 7.5, before addition of 200 µM of GlcN-1P and Ac-CoA to initiate the reaction and 400 µM DTP to monitor formation of CoA-SH, which was monitored kinetically (y-axes) (A) Titrations of the two WT preparations of GlmU; HT purified enzyme (●) and GlmU^{+Mg2+} (●), compared to GlmU H374A (●). (B) Titrations of the mutant GlmU preparations; R344A (●), K362A (●) and Y377A (●). Data points are an average of three replicates and solid lines are the fit of the data to a linear regression.

The HT preparation of GlmU WT had equivalent, if not superior, acetyltransferase activity to the previous preparation of WT enzyme, GlmU^{+Mg2+}, which was purified using an Äkta column chromatographic process. As had previous been observed, the GlmU H374A mutant was inactive, in the pH and enzyme concentration range assessed in this experiment. The GlmU SDM mutant proteins; R344A, K362A and Y377A, were significantly less catalytically active than the WT enzyme, by between 11 to 25-fold at equivalent concentrations. GlmU Y377A is catalytically inactive at neutral pH and within the concentration range utilised for this experiment. These data demonstrated that the Tyr377 to Ala mutation, had a similar effect to the His374 to Ala change, which was the ablation of acetyl transfer activity when compared to the WT protein in assay buffer adjusted to pH 7.5. The data presented in Figure 83 (B) demonstrated that the mutation of either Arg344 or Lys362 to Ala, while not null mutations in the assay conditions utilised, still significantly reduce GlmU

acetyltransferase activity. Further studies will be carried out to understand whether varying the concentration of the acetyltransferase substrates or pH conditions of the assay buffer have any effect on the acetyl transfer activity of either GlmU R344A, K362A or Y377A.

4.6 Further GlmU SDM pH-rate profiles

To determine the potential involvement of Arg344, Lys362 or Tyr377 in the GlmU-catalysed acetyl transfer reaction, pH-rate studies were undertaken by monitoring acetyltransferase activity at varying Ac-CoA and GlcN-1P, while utilising saturating second substrate concentrations, in assay buffer adjusted to different pH values. The data presented in Figure 84 was generated by titrating GlcN-1P in varying pH conditions across the range of 6.0 to 10.0, in the presence of saturating concentrations of Ac-CoA at each condition.

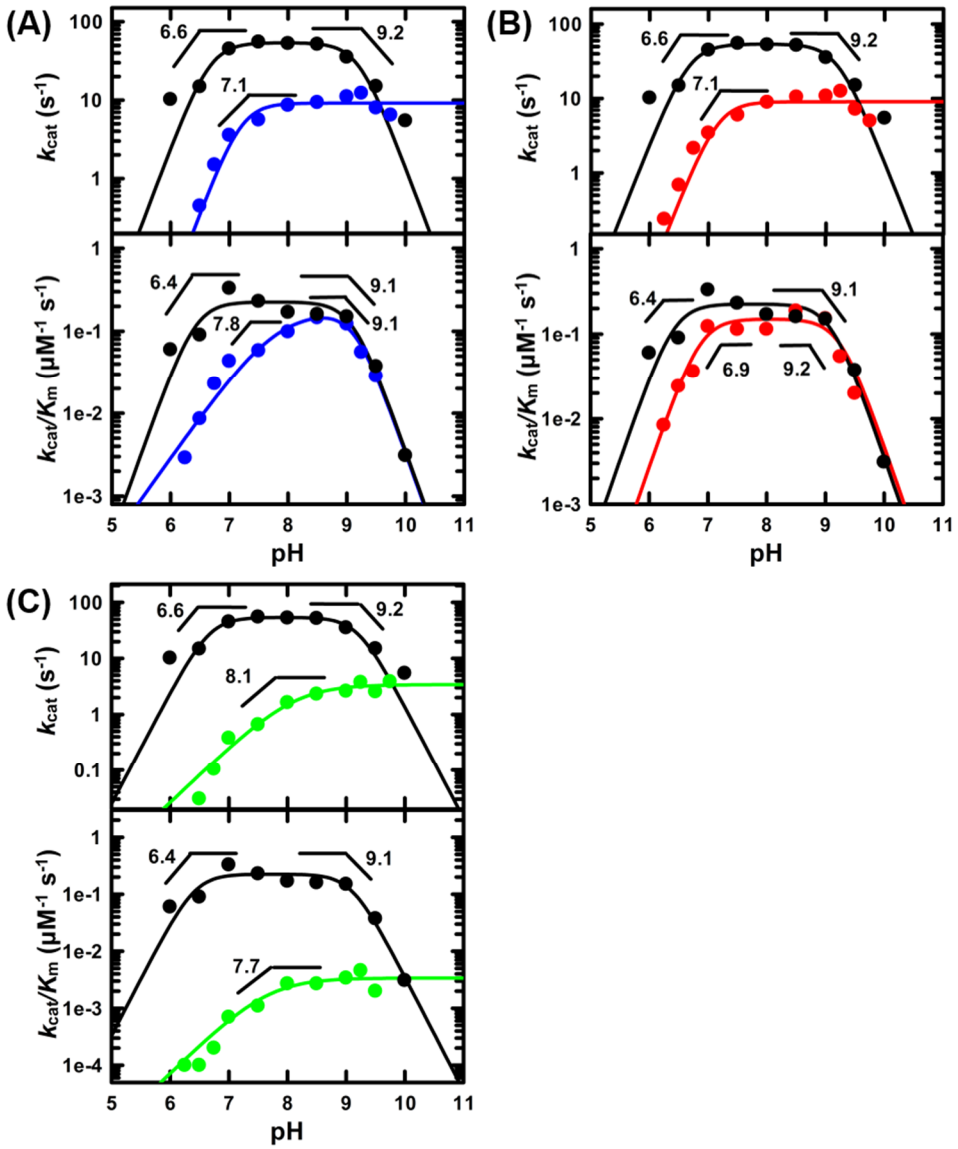


Figure 84. Acetyltransferase pH-rate profiles of GlmU WT and SDM mutants Ac-CoA kinetic parameters.

GlmU acetyltransferase activity k_{cat} and K_m values were determined in a range of pH conditions by varying concentration of Ac-CoA at fixed, saturating concentrations of GlcN-1P. In all plots, GlmU WT data is represented by (●). The GlmU WT k_{cat} and k_{cat}/K_m data were fitted to Eqn. (11) and are represented by (—). (A) The GlmU R344A (●) k_{cat} and k_{cat}/K_m data were fitted to Eqn. (10) and Eqn. (14), respectively, and are represented by (—). (B) The GlmU K362A (●) k_{cat} and k_{cat}/K_m data were fitted to Eqn. (10) and Eqn. (11), respectively, and are represented by (—). (C) The GlmU Y377A (●) k_{cat} and k_{cat}/K_m data were both fitted to Eqn. (9) and are represented by (—).

Table 21. GlmU WT and SDM mutant acetyltransferase pH profile ionisable group pK_a values where Ac-CoA is the varied substrate.

Construct	k_{cat} pH profile acidic pK_a [a, b]	k_{cat} pH profile basic pK_a [a, b]	k_{cat}/K_m pH profile acidic pK_a [a, b]	k_{cat}/K_m pH profile basic pK_a [a, b]	k_{cat} pH profile equation used [c]	k_{cat}/K_m pH profile equation used [c]
WT	6.66 ± 0.07	9.20 ± 0.07	6.42 ± 0.24	9.09 ± 0.24	(11)	(11)
R344A	-	7.20 ± 0.15	7.78 ± 0.24	9.14 ± 0.06	(10)	(14)
K362A	-	7.17 ± 0.19	6.87 ± 0.14	9.25 ± 0.14	(10)	(11)
Y377A	-	8.11 ± 0.17	-	7.66 ± 0.27	(9)	(9)

[a] Each data set is a mean of three replicates. [b] Methodology is detailed in Section 2.2.20. [b] Data analysis details can be found in Section 2.2.30.

The GlmU SDM mutants are slower acetyltransferase enzymes than GlmU WT, displaying a pH-independent k_{cat} plateau that is either 5-fold (for R344A and K362A) or 25-fold (Y377A) lower (Figure 84). The pH dependence of both GlmU R344A and K362A k_{cat} reveals two non-resolvable ionisable groups that must be deprotonated for maximal activity (slope of +2). These data were both best fitted to Eqn. (10), see Table 21 for pK_a values. The pH dependence of GlmU Y377A k_{cat} reveals one ionisable groups that must be deprotonated for maximal activity (slope of +1), which was best fitted to Eqn. (9).

The plot of pH versus GlmU R344A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ showed both a single ionisable acidic (slope of +1) and two non-resolvable basic groups (slope of +2), which was best fitted to Eqn. (14). The $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH plot for GlmU K362A identified two non-resolvable acidic and basic groups, which were best fitted to Eqn. (11). GlmU Y377A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH plot identified one basic group, which was best fitted to Eqn. (9).

Further acetyltransferase pH-rate profile data for GlmU WT was generated in parallel with the GlmU SDM study. This control study, utilised Ac-CoA as the varied substrate in the presence of saturating GlcN-1P, was best fitted to the equation for two nonresolvable acidic and basic groups (Eqn. (12)). The pKa values for the ionisable groups are presented in Table 21.

In addition to determining the GlmU SDM catalysed Ac-CoA Michaelis-Menten parameters in varying pH conditions, the GlcN-1P steady state parameters were determined, utilising the same pH range, along with WT controls (Figure 85).

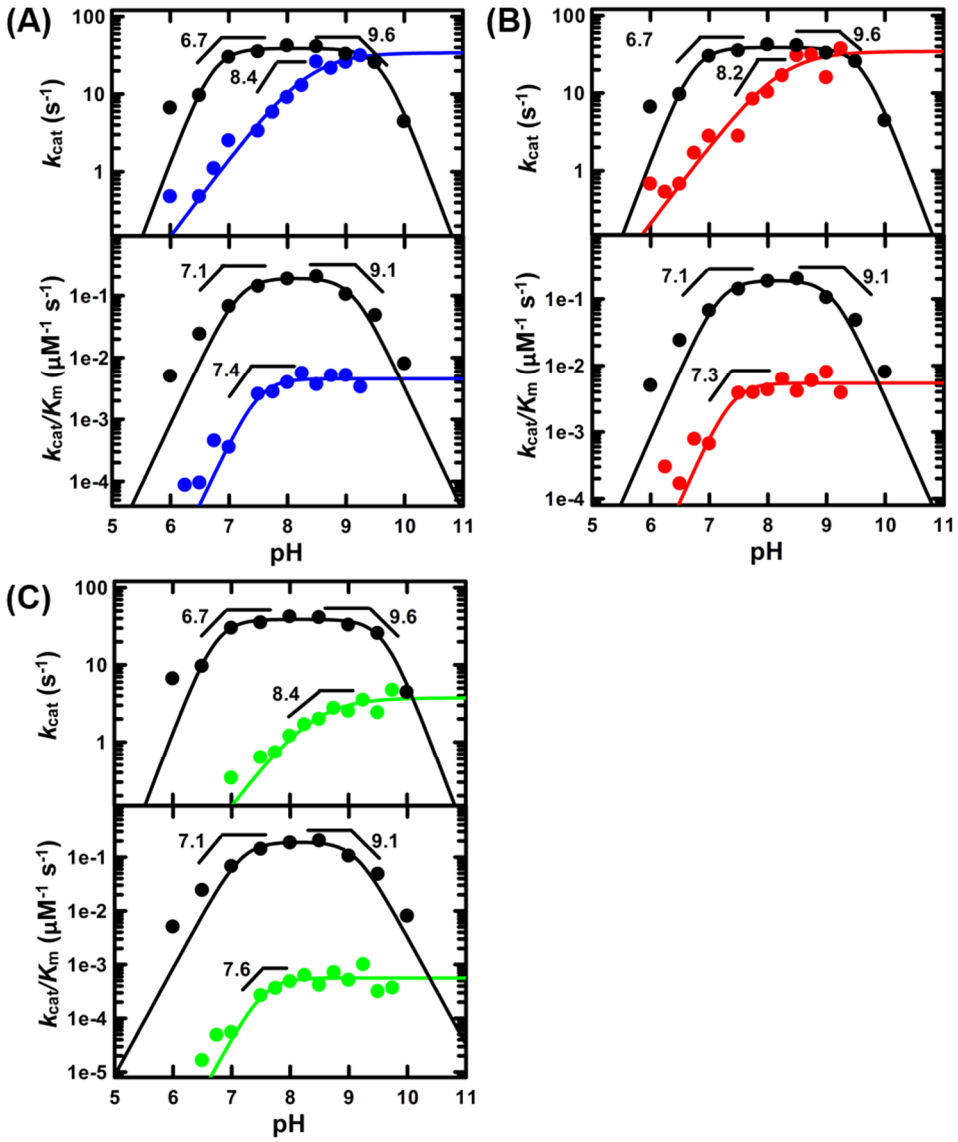


Figure 85. Acetyltransferase pH-rate profiles of GlmU WT and SDM mutants GlcN-1P kinetic parameters.

GlmU acetyltransferase activity k_{cat} and K_m values were determined in a range of pH conditions by varying concentration of GlcN-1P at fixed, saturating concentrations of Ac-CoA. In all plots, GlmU WT data is represented by (●). The GlmU WT k_{cat} and k_{cat}/K_m data were fitted to Eqn. (11) and are represented by (—). (A) The GlmU R344A (●) k_{cat} and k_{cat}/K_m data were fitted to Eqn. (13) and Eqn. (10), respectively, and are represented by (—). (B) The GlmU K362A (●) k_{cat} and k_{cat}/K_m data were fitted to Eqn. (13) and Eqn. (10), respectively, and are represented by (—). (C) The

GlmU Y377A (●) k_{cat} and k_{cat}/K_m data were fitted to Eqn. (13) and Eqn. (10), respectively, and are represented by (—).

Table 22. GlmU WT and SDM mutant acetyltransferase pH profile ionisable group pK_a values where GlcN-1P is the varied substrate.

Construct	k_{cat} pH profile Acidic pK_a [a, b]	k_{cat} pH profile Basic pK_a [a, b]	k_{cat}/K_m pH profile Acidic pK_a [a, b]	k_{cat}/K_m pH profile Basic pK_a [a, b]	k_{cat} pH profile equation [c]	k_{cat}/K_m pH profile equation [c]
WT	6.73 ± 0.08	9.61 ± 0.08	7.18 ± 0.08	9.11 ± 0.08	(11)	(11)
R344A	-	8.37 ± 0.11	-	7.51 ± 0.11	(13)	(10)
K362A	-	8.22 ± 0.22	-	7.39 ± 0.17	(13)	(10)
Y377A	-	8.40 ± 0.16	-	7.55 ± 0.21	(13)	(10)

[a] Each data set is a mean of three replicates. [b] Methodology is detailed in Section 2.2.20. [b] Data analysis details can be found in Section 2.2.30.

The data presented in the pH profiles of the GlmU mutants R344A and K362A in the k_{cat} plot of Figure 85, is markedly different to the data in the k_{cat} plot of Figure 84. The data in Figure 85, which determined the pH dependence of the acetyltransferase kinetic parameters for GlcN-1P, demonstrated that the same magnitude of $k_{cat,\ GlcN\text{-}1P}$ as the WT enzyme can be achieved with the R344A and K362A mutants, at pH 9 to 9.5. This observation was significantly influenced by a 10-fold reduction in $K_{m,\ GlcN\text{-}1P}$ for the SDM mutants, which lead to non-saturating concentrations of GlcN-1P being utilised for the Ac-CoA pH-rate studies (Figure 84). Experimentally it was not possible to saturate with the second substrate when titrating Ac-CoA. The same issue is true for both the GlmU Y377A Ac-CoA and GlcN-1P pH-rate profile studies. The Tyr377 to Ala mutation has the effect of reducing $K_{m,\ Ac\text{-}CoA}$ by 10-fold, in comparison to the WT and R344A and K362A mutants. In addition, the Tyr377 to Ala mutation also reduces $K_{m,\ GlcN\text{-}1P}$ by 10-fold, in comparison to the WT, consequently, sub-saturating concentrations of second substrate were used in both the Ac-CoA and GlcN-1P GlmU Y377A pH-rate profile studies. The apparent pK_a data has been presented as it is insightful and will potentially allow further elucidation

of the roles of key active site residues involved in acetyltransferase substrate binding and catalysis.

The pH dependence of both GlmU R344A and K362A $k_{cat,\ GlcN\text{-}1P}$ reveals one ionisable groups that must be deprotonated for maximal activity (slope of +1), which were both best fitted to Eqn. (13). The pH dependence of GlmU Y377A $k_{cat,\ GlcN\text{-}1P}$ reveals one ionisable groups that must be deprotonated for maximal activity (slope of +1), which were best fitted to Eqn. (13).

The plot of pH versus GlmU R344A $k_{cat}/K_{m,\ GlcN\text{-}1P}$ showed two non-resolvable ionisable acidic groups (slope of +2), which were best fitted to Eqn. (10). The $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH plot for GlmU K362A identified two non-resolvable acidic groups, which were best fitted to Eqn. (10). GlmU Y377A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH plot identified two non-resolvable acidic groups, which were best fitted to Eqn. (10).

Further acetyltransferase pH-rate profile data for GlmU WT was generated in parallel with the GlmU SDM study. This control study, utilised GlcN-1P as the varied substrate in the presence of saturating Ac-CoA, was best fitted to the equation for two nonresolvable acidic and basic groups (Eqn.(11)).

These data, taken collectively, allow a greater understanding of the role of the investigated amino acids in both acetyltransferase substrate recognition, binding and catalysis. However, these studies are not sufficient to assign all the measured WT p*K* values and further work is required to ascertain the identifies of all titratable groups involved in substrate binding and GlmU catalysed acetyl transfer.

4.7 Acetyltransferase solvent kinetic isotope effect

Solvent kinetic isotope effects (SKIEs) were utilised to investigate the acetyl transfer chemical mechanism. Proton transfers are essential in this mechanism and therefore SKIEs might identify solvent-derived proton transfer events that might be important during catalysis. To carry out these solvent kinetic isotope studies, buffers and assay reagents, including substrates and the DTP CoA-SH detection reagents were prepared in either 100% H_2O or 100% D_2O. The only reagent not to be prepared in D_2O was GlmU, however, the enzyme is used at 5 nM in the acetyltransferase absorbance assay and the dilution from stock is such that there is only 0.05% H_2O contamination in the D_2O experiments. D_2O is more viscous than

H_2O and this difference in viscosity must be controlled for, either prior to or in parallel to the SKIEs studies. The viscosity control that is used as a matter of course for these type of studies is to replicate the saturation titrations of each of the substrates in H_2O supplemented with 9% glycerol, which acts as a viscogen, and closely mimics the increased viscosity of D_2O. To ensure that the maximal velocity was utilised for these studies, pL 8 was selected as it corresponds to centre point of the plateau region of the k_{cat} pH-rate profile for the WT enzyme, where the variation due to any changes in experimental pH/pD are smallest.

To carry out the SKIEs studies either GlcN-1P or Ac-CoA were titrated from a high concentration in the presence of fixed, saturating concentrations of the corresponding second substrate, prior to addition of an assay initiation mix including, 5 nM GlmU, to initiate acetyltransferase activity, and DTP, to kinetically monitor of CoA-SH formation by a change in absorbance at 324 nm, over a 30 mins period. These experiments were replicated utilising either reagents prepared entirely in D_2O or H_2O supplemented in 9% glycerol (Figure 86).

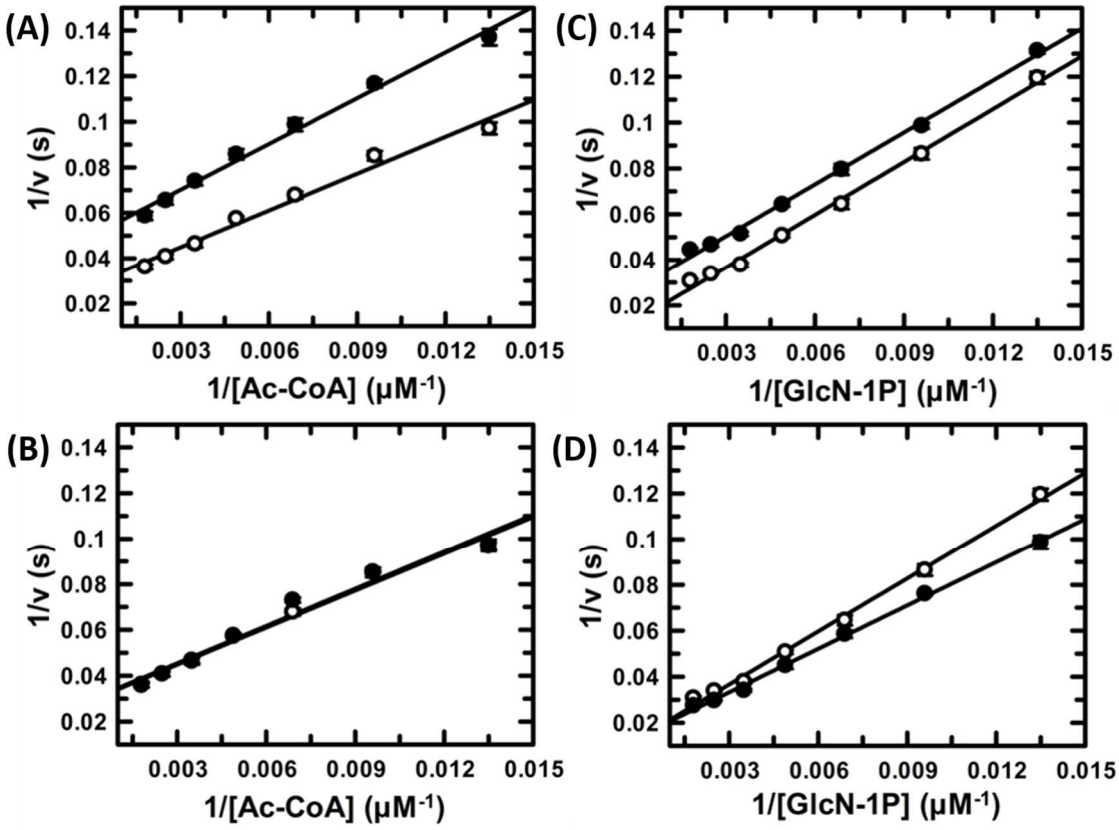


Figure 86. GlmU acetyltransferase SKIEs and viscosity controls.

Panels (A) and (B); Ac-CoA was used as the varied substrate, with a saturating concentration of GlcN-1P. (A) The reaction mix contained either 0 (○) or 99.95 (●) atom % D_2O. (B) The reaction mix was prepared in H_2O supplemented with either 0 (○) or 9 (●) % glycerol. Panels (C) and (D); GlcN-1P was used as the varied substrate, with a saturating concentration of Ac-CoA. (C) The reaction mix contained either 0 (○) or 99.95 (●) atom % D_2O. (D) The reaction mix was prepared in H_2O supplemented with either 0 (○) or 9 (●) % glycerol. The experimental data, which in all cases were a mean of six replicates, are represented by circles and the best fit of data are represented by ─. (A) data fitted to Eqn. (17), (B) data fitted to Eqn. (17), (C) data fitted to Eqn. (15), and (D) data to fitted to Eqn. (16).

The GlmU acetyltransferase Ac-CoA and GlcN-1P Michaelis-Menten parameters determined for the solvent kinetic isotope effect and viscosity control data presented in both Figure 86 can be found in Table 23.

Table 23. GlmU acetyltransferase SKIEs and viscosity control Michaelis-Menten data [a].

varied substrate	fixed substrate	^{D_2O}V	$^{D_2O}V/K$
solvent kinetic isotope effect			
Ac-CoA	GlcN-1P	1.61± 0.03	1.36 ± 0.08
		1.61 ± 0.03 [b]	1.32 ± 0.1 [b]
GlcN-1P	Ac-CoA	1.59 ± 0.07	1
		1.55 ± 0.07 [b]	1
viscosity effect [c]			
Ac-CoA	GlcN-1P	1	1
GlcN-1P	Ac-CoA	1	0.76 ± 0.08

[a] Solvent isotope effects were determined in assay buffer prepared in either 100% H_2O or 100% D_2O, at pH 8 and 30 °C. [b] Solvent isotope effect viscosity control using 9 % glycerol in assay buffer as a viscogen.

The data in Table 23 shows that there is a clear SKIE on *V* when GlcN-1P is varied and a SKIE on both *V* and *V/K* when Ac-CoA was varied.

To determine the number of protons that are in flight during the transition-state giving rise to the measured SKIEs, a proton inventory study was carried out. This involves

measuring the velocity of the acetyltransferase reaction, at pL 8, utilising saturating concentrations of GlcN-1P and Ac-CoA, as a function of the atom fraction of deuterium in a mixed H_2O/D_2O solvent system (Figure 87).

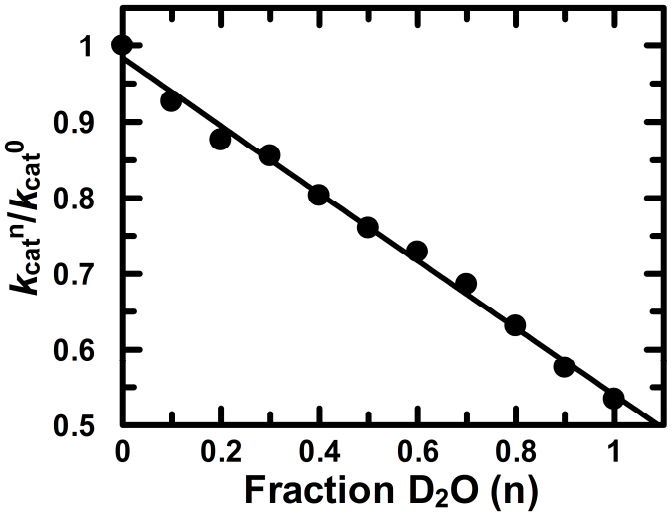


Figure 87. GlmU acetyltransferase proton inventory.

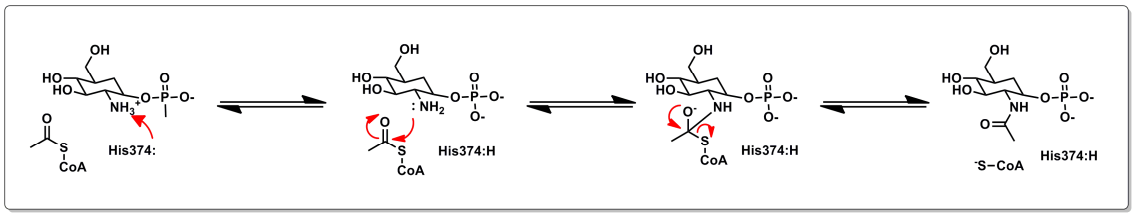
Acetyltransferase activity monitored by formation of CoA-SH in assay buffer prepared in H_2O, and in the presence of 5 nM GlmU, 1.5 µM GlcN-1P and 1 µM Ac-CoA. These conditions were repeated at varying fractions of D_2O up to and including 100 % D_2O. The y axis is the ratio of k_{cat} in a mixture of D_2O and H_2O (k_{cat}^n) and the k_{cat} value in 100% H_2O (k_{cat}^0). The data are a mean of at least 15 replicates and were subsequently fit to Eqn. (18).

The data in Figure 87 was best fit to Eqn. (18), which describes a linear function that is diagnostic of a single proton transfer event being responsible for the solvent isotope effect.

These data demonstrated that there are modest SKIEs on V and $V/K_{Ac\text{-}CoA}$ indicating that acetyl transfer is partially rate-limiting for the overall reaction. Further studies, utilising a proton inventory experiment, identified that the observed SKIE on V is due to a single proton in flight.

4.8 Proposed chemical mechanism of GlmU acetyltransferase activity

The results presented in Chapter 4, provide evidence for the proposal of a chemical mechanism for the GlmU catalysed acetyltransferase activity, which is presented in Scheme 13.



Scheme 13. Proposed chemical mechanism for *M. tuberculosis* GlmU acetyltransferase activity

The proposed chemical mechanism in Scheme 13 employs general base catalysis, which is supported by the observations from pH-rate profiles, which identified a residue with a pK_a of ~ 6.5 required for catalysis. The identity of the group is likely to be His374, which acts as a general base

Additionally, the results presented within this chapter begin to provide evidence for the elucidation of additional residues that are important for GlmU catalysed acetyl transfer to GlcN-1P. SDM of likely important acetyltransferase active site residues, followed by pH-rate profile studies identified that Arg344 and Lys362 are important for the binding of GlcN-1P and mutation of these groups to Ala led to a change in the pH rate profiles, in comparison to wild-type GlmU. The mutation of Tyr377 to Ala also led to a change in the pH-rate profiles, when compared to wild-type, R344A and K362A. These data suggest that Tyr377 has not only an effect on substrate binding but also on catalysis, as the fold-change in both k_{cat} and k_{cat}/K_m pH-rate profiles, when compared to the wild-type enzyme, is greater than that of the R344A and K362A mutants.

Finally, the SKIE studies identified that one proton is in flight during catalysis and that the substitution of H_2O for D_2O, predominantly has an effect on k_{cat}, suggesting that the partially rate-limiting step is due to either catalysis or product release.

The data presented in Chapter 4 will be discussed in more depth in the discussion section of this thesis, Chapter 6.

Chapter 5. GlmU acetyltransferase small molecule inhibitor identification and hit qualification

In parallel to, and informed by, the acetyltransferase enzyme mechanism studies presented in the previous two chapters, GlmU small molecule inhibitor identification studies were carried out. The aims of these studies were two-fold, initially, to identify novel chemical compounds that could be used to further validate the acetyltransferase mechanism of GlmU, and because of these efforts, provide evidence to validate this bifunctional enzyme as a viable target for more thorough drug discovery efforts.

Classically, for small molecule screening studies, several assay detection approaches are evaluated to validate their use with the test protein of interest. The most important factor typically assessed is the sensitivity of the detection technique to the substrate or product that will be monitored. The sensitivity is typically defined by the minimum concentration of either substrate or product that can be detected by an assay methodology. This has already been addressed for GlmU, to some degree in Section 3.2 but will be further expanded upon in this chapter.

Small molecule inhibitor identification requires consideration of a number processes of scale that enable many targeted and diverse compounds to be tested against the drug target of interest. These considerations include the amenability of the primary assay to miniaturisation, automation of reagent addition, reproducibility of the assay, stability of reagents over prolonged periods at 4 °C, and tolerance of the assay to solvent (dimethyl sulphoxide (DMSO), commonly used to dissolve test compounds). These considerations have been systematically addressed for the selected GlmU acetyltransferase assay prior to commencement of inhibitor screening.

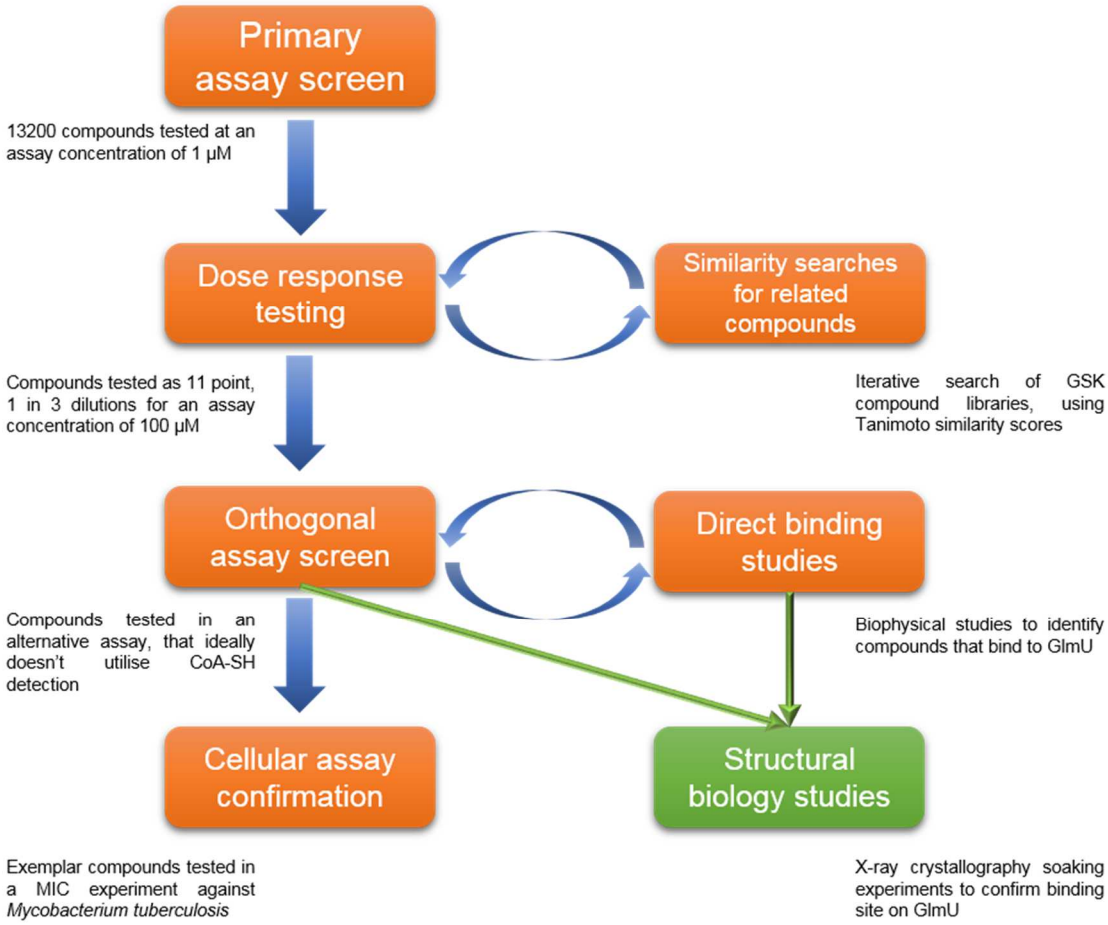
The suitability or quality of the assays selected for acetyltransferase inhibitor identification screening were defined by evaluating a screening window coefficient, known as the Z' factor, which is a combined measure of the assay signal dynamic range and the variation associated with the assay signal variation.

Prior to progressing with single concentration, inhibitor identification screening, two training sets were screened to define both the likely hit rate and identify potential, compound mediated assay interference. Hit rate, in this context, is the number of potential inhibitors that are statistically separated from the inactive compound

population. The first small molecule training set, known as the “Robustness set”, which is a collection of 1408 diverse compounds representative of the GSK HTS collection, is typically screened at a single concentration, to define the inhibitor identification, performance of the primary assay of choice.

The second small molecule set, known as the “Nuisance set”, was put together to identify the susceptibility of the screening assay to a group of compounds known as PAINS (Section 1.4.1)[426]. The outcome of screening this set identifies the type of interferences that will appear as inhibitors in the assay, commonly known as false positives, and provide knowledge to remove these false positives after the initial hit identification screening.

The small molecule sets screened to identify potential inhibitors of GlmU acetyltransferase activity were comprised of two sets made up of compounds that have either been shown to kill or inhibit *Mycobacterium tuberculosis*[427]. The third set that was screened was a diverse collection of small molecules, known as the “Validation set”, made up of 10000 compounds, which are representative, of the GSK HTS collection, covering a wider selection of chemical diversity than the Robustness set. The total number of small molecules that were screened against GlmU acetyltransferase activity, as a single compound concentration, in the hit identification studies was 13200. To confirm the identified hit compounds are more likely to be assay methodology independent inhibitors, two approaches were undertaken. The first approach is to test the hits at multiple concentrations, which will confirm that the acetyltransferase activity inhibition is compound dose dependent. The second approach is to test the potential hits in both an orthogonal acetyltransferase assay, as well as investigating direct compound binding using a biophysical technique. This inhibitor identification strategy, which is typical of a number of drug discovery approaches employed by pharmaceutical and biotechnology companies[428], is best represented by Scheme 14.



Scheme 14. GlmU acetyltransferase inhibitor identification strategy.

Confirmation of hit compounds was carried by utilisation of an orthogonal assay to monitor GlmU acetyltransferase activity. To best utilise the benefits of this approach it was essential to select an assay that was not dependent on absorbance and the detection of the formation of CoA-SH to monitor activity. The approach developed and utilised was a high throughput MS method, known as Agilent RapidFire (RF-MS), which employs a rapid microplate sampling probe to aspirate an aliquot of the stopped assay, prior to sample purification on a solid phase extraction cartridge (to remove problematic contaminants) before injection into the MS[429]. This method was used to monitor the relative concentrations of GlcN-1P and GlcNAc-1P, which could be modulated by a small molecule inhibitor. This assay was used to confirm acetyltransferase product inhibition by CoA-SH in Section 3.11 and further details regarding the development of this approach, as well as application to confirm GlmU acetyltransferase inhibitors will be presented in this chapter.

Biophysical studies were utilised to determine if potential inhibitory small molecules directly bound to GlmU, which provided further confirmatory evidence for inhibition by these compounds. DSF has been shown to be an insightful and reliable approach for monitoring direct binding of substrates, products and analogues to GlmU and was further utilised to confirm interaction of potential inhibitors with the enzyme by stabilising the $T_{m.}$ The application of DSF to identify and confirm small molecule binding partners is well documented[407] and used as a matter of routine by a number of companies and research groups[300,430].

As a parallel approach, knowledge-based, rational inhibitor design was carried out to explore whether synthetic analogues of GlcN-1P could act as antagonists of the GlmU acetyltransferase activity. Rational inhibitor design using one of the native substrates, products or transition states as starting points is a well-accepted approach to develop novel antagonists[301]. As with previous studies to identify dead-end inhibitors of GlmU acetyltransferase (Section 3.11), these synthetic analogues were first investigated as substrates, as a control before investigating whether any of these molecules were indeed inhibitors.

To confirm inhibition of small molecule inhibitors, exemplars were screened against *Mycobacterium tuberculosis* H37Rv as per the methodology detailed in Section 2.2.28 and by Ballell and co-workers[427]. This assay defines the MIC for *Mycobacterium tuberculosis* of the test compounds and confirms if the novel inhibitors have the potential to be antimicrobial agents.

The findings of these initial small molecule identification studies will help define whether the acetyltransferase domain of GlmU is a tractable target for further, more concerted and better resourced hit identification efforts. In addition, any inhibitors identified and validated in this study could provide either useful starting points for further chemical optimisation, in addition to being powerful tools to better characterise the essentiality of GlmU for *Mycobacterium tuberculosis* growth and survival.

All equations used in the analysis of data in this chapter can be found in the Materials and Methods chapter, section 2.2.30.

All compounds tested in this section, unless otherwise stated, are GSK intellectual property and cannot be disclosed without a specific request, which can be made via the author of this thesis or their supervisor. Where GSK compound registration numbers are clearly stated then the structures of these molecules have previously

been disclosed most likely in manuscripts authored by either Ballell, *et al.* or Rebollo-Lopez, *et al.*[335,336].

5.1 Compound screening assay configuration

GlmU acetyltransferase assay identification for the mechanistic studies detailed in Chapters 3 and 4, has previously been detailed in Section 3.2, where the DTP thiol detection absorbance assay was described, however, as there are several additional considerations required for the configuration of an accurate and sensitive small molecule screening assay.

Sensitivity of the detection methodology utilised has previous been mentioned in preceding Chapter introduction and is important to ascertain early in assay development. The data presented in Section 3.5, determines the GlmU acetyltransferase Michaelis-Menten parameters as $K_{m,\ Ac\text{-}CoA}$ = 240 µM and $K_{m,\ GlcN\text{-}1P}$ = 344 µM and subsequently concentrations of 200 µM of these substrates were used to assess dead-end substrate analogues in Section 3.11. The use of substrate concentrations equivalent to the apparent K_m value or below is to ensure that, when evaluating potential inhibitors, the assay is not biased towards a substrate bound form, in addition, the assay is linear with relation to substrate concentration (Figure 88).

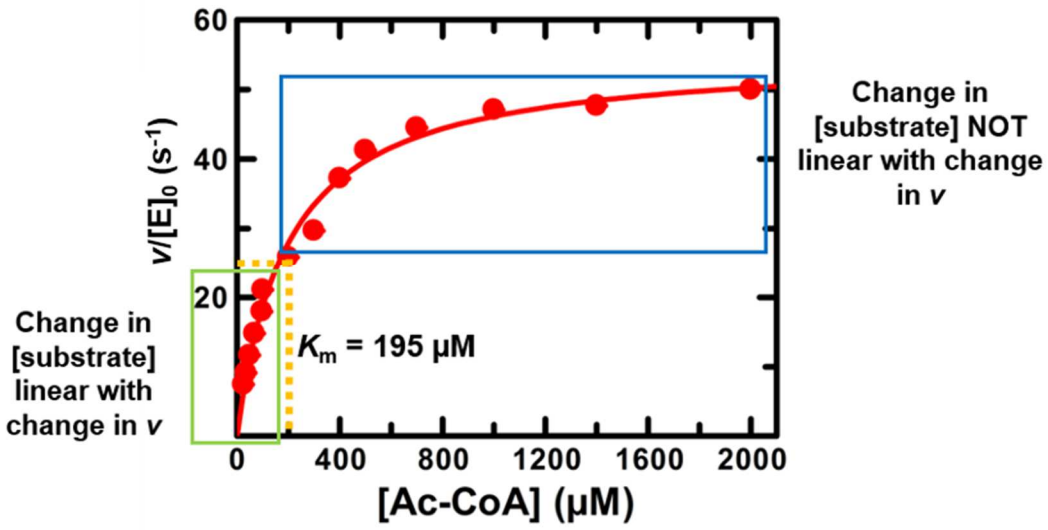


Figure 88. Configuration of GlmU acetyltransferase DTP absorbance screening assay at $K_{m,\ Ac\text{-}CoA}$.

Titration of Ac-CoA from 2 mM in the presence of 2 mM GlcN-1P in assay buffer adjusted to pH 7.5. Acetyltransferase activity was initiated by addition of 5 nM GlmU and 400 µM DTP, which reacts with CoA-SH, leading to a change in absorbance at 324 nm. The K_m, Ac-CoA is highlighted, as are the concentration ranges where the change in Ac-CoA is linear with *v*, in addition to range where the change in substrate is not linear with *v*.

The representative data presented in Figure 88 highlights the ideal concentration range of Ac-CoA for a sensitive acetyltransferase assay, where a 50% decrease in substrate will lead to the equivalent reduction in measured *v*. Data for GlcN-1P is not presented but the same ranges are true for this substrate. It should be noted that configuration of the GlmU acetyltransferase compound screening assay at substrate concentrations equivalent to the apparent K_m values ensured that all modes of inhibition can be detected (competitive, noncompetitive and uncompetitive). Configuration of the absorbance assay at concentrations of substrates significantly more than the K_m values would bias detection toward noncompetitive and uncompetitive inhibition and make detection of competitive inhibitors less likely.

The concentration of DTP utilised to detect CoA-SH had to be more than the maximum concentration of Ac-CoA added to the assay to prevent depletion of the thiol detection reagent. DTP does not absorb light at 324 nm if there is not sufficient concentration of thiol to react with, so adding an excess of detection reagent is not detrimental to configuration of the assay. The data presented in Figure 26 demonstrated that the detection of thiols by DTP is linear between 1 µM or 500 µM. Further consideration must be given to the concentration of GlmU used, which is linked to several factors including; the potency of inhibitors that can be detected, substrate turnover and practically, kinetic read duration. To illustrate the interplay between concentration of GlmU used and the percentage of Ac-CoA catalysed, linear progression plots of a limited range of enzyme concentration are presented in Figure 89.

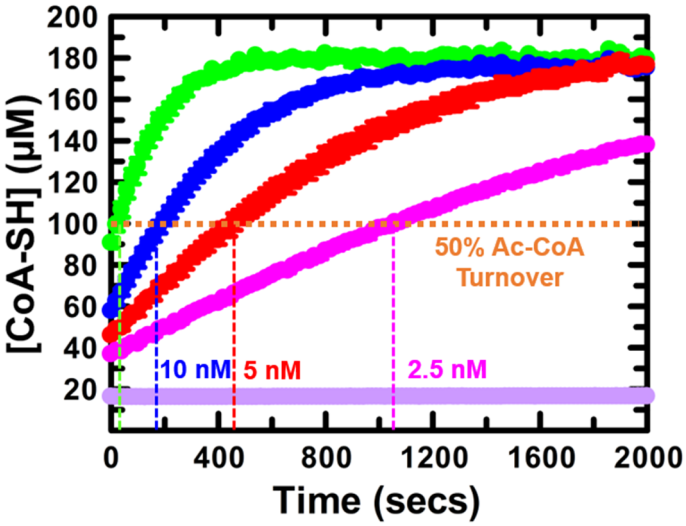


Figure 89. GlmU acetyltransferase activity progress plots.

GlmU acetyltransferase activity in the presence of 200 µM Ac-CoA and GlcN-1P, in assay buffer adjusted to pH 7.5, utilising 400 µM to detect formation of CoA-SH. The progress plots presented are for the following concentrations of GlmU; 20 nM (●), 10 nM (●), 5 nM (●) and 2.5 nM (●). The no enzyme background is represented by (●). The data points represent a mean of three independent replicates, error bars added. The level of 50% Ac-CoA conversion to CoA-SH is highlighted by the dashed orange line.

The data presented in Figure 89 demonstrates that when 5 nM GlmU is utilised, 50% conversion of Ac-CoA to CoA-SH is achieved within 8 mins, whereas, the same level of product formation is attained by 20 nM GlmU after 1 min. Finally, the 2.5 nM enzyme concentration catalyses formation of 100 µM CoA-SH in 18 mins. The minimisation of substrate conversion is key, as substrate depletion can lead to desensitisation of the assay to inhibition by test small molecules[431]. The concentration of 5 nM GlmU was selected for further screening studies, as using this condition the acetyltransferase activity was linear for sufficient time to screen the test microplates. Additionally, the difference between the 5 nM GlmU acetyltransferase response and the background signal, in the absence of enzyme, when taking into account the standard deviation of both populations was sufficient to generate a passable Z' factor in excess of 0.4 (Figure 90).

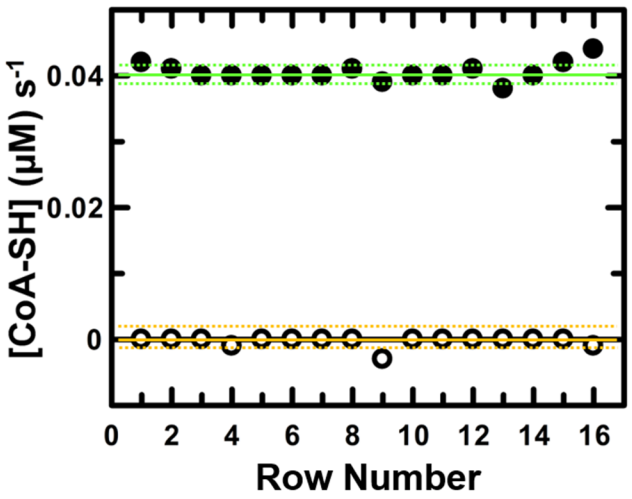


Figure 90. Z' factor of GlmU acetyltransferase absorbance assay.

GlmU acetyltransferase assay signal window demonstrating a population with 5 nM enzyme (●) and no enzyme control (○). The mean of the two populations are displayed as either plus enzyme (—) or minus enzyme (—). The standard deviations of the two populations are displayed as either plus enzyme (---) or minus enzyme (---).

The minimum acceptable Z' value for an enzyme assay of this type at GSK is 0.4, while a good assay would be in the range of 0.6 to 0.9. The data in Figure 90 when fitted to Eqn. (24) yielded a Z' = 0.84, which demonstrated that the GlmU acetyltransferase DTP absorbance assay was of a high quality (the maximum Z' value is 1). It should be noted that the data presented in Figure 90 was generated utilising liquid handling automation, which validated the use of these instruments for further higher throughput screening studies. The significance of a good Z' factor is that there is a greater confidence in the hit compounds identified during single concentration small molecule screening, as there is a greater separation and less variation, between the inhibited and non-inhibited values[399]. This Z' value was consistent across a number of initial test occasions and further data will be presented in the coming sections of this chapter.

The final assay protocol for the GlmU acetyltransferase DTP absorbance screening assay is detailed in Section 2.2.22.

Initial configuration of the acetyltransferase screening assay was completed, and further studies will be carried out with single concentration compound training sets.

5.2 Compound training set screening

The initial single concentration, small molecule training set screened was the GSK Robustness set, which is comprised of 1408 diverse compounds. These compounds are dispensed onto 4, 384 well microplates and are typically screened in duplicate on two separate test occasions to ascertain the reproducibility of the assay. The data presented in Figure 91 is the correlation of the two test occasions of GlmU acetyltransferase absorbance assay.

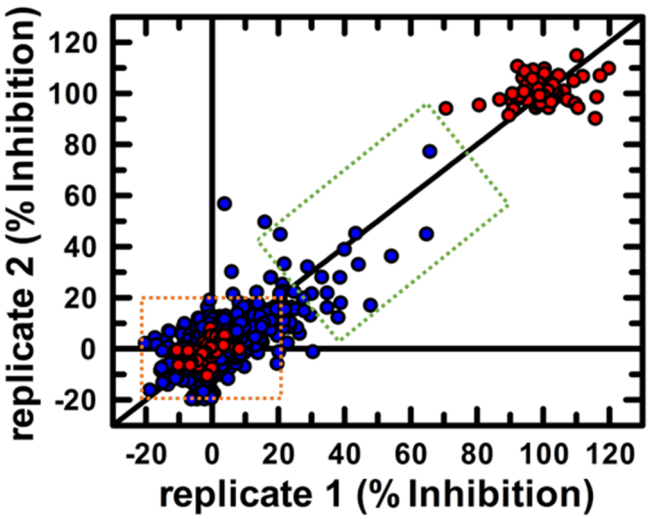


Figure 91. GlmU acetyltransferase assay robustness correlation.

GSK robustness set screened on two separate test occasions and the data normalised to the controls on each of the test microplates using Eqn. (23). The positive and negative control populations are highlighted by (●) and the test wells are identified by (●). The orange box highlights the statistical cut off that determines whether a test compound is identified as active or inactive. The green box highlights the compounds that were selected for dose response testing.

The GlmU acetyltransferase activity Robustness screen identified 21 compounds that were statistically classed as hits on both test occasions, which as a percentage of the 1408 compounds tested is 1.5%. The statistical threshold used to classify compounds as either a hit or inactive is demonstrated in Figure 92.

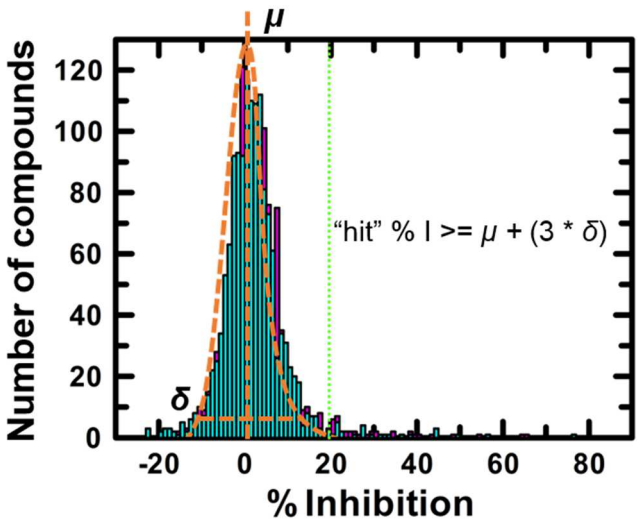


Figure 92. Distribution of % Inhibition for robustness set test compounds.

Histograms of the distribution of % Inhibition values of the 1408 test compounds of the GSK robustness set when screened in the GlmU acetyltransferase assay. The histograms in teal and purple are the % Inhibition data for replicate 1 and 2, respectively. The dashed orange lines highlight both the normal nature of the % Inhibition distribution and the mean (μ) and the standard deviation (δ) of the test compounds. The hit threshold is highlighted by the dashed green line and is calculated by using the mean and standard deviation of the test samples.

To classify a test compound as a hit, robust statistics were used to determine the mean and standard deviation of the samples, which typically form a normal distribution as seen in Figure 92. This distribution is skewed as the sample population comprises both hit and inactive compounds, so robust statistics are utilised to weight out the hits and enable determination of a non-skewed, robust mean and standard deviation[432]. A test compound is classified as a hit if the % Inhibition value measured in the GlmU acetyltransferase assay is above a threshold that is determined by multiplying the robust standard deviation of the % Inhibition of the samples by 3 and then adding this to robust mean of the % Inhibition values of all the test compounds[433]. The threshold for identifying a hit compound determined for the two robustness set test occasions was 19.5 % Inhibition.

The Z' factor was determined from two control populations of 16 per microplate tested, which were uninhibited (0% Inhibition) and a no-enzyme controls (100% Inhibition). The Z' factor and the microplate controls are presented in Figure 93.

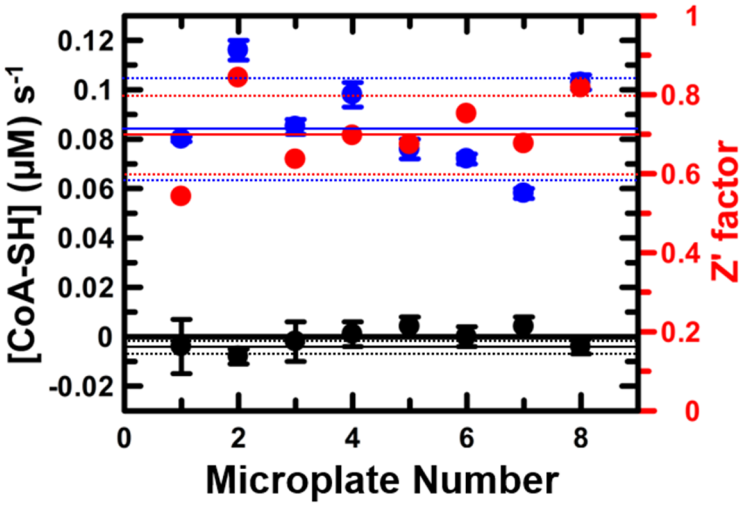


Figure 93. Z' factor and assay controls of GlmU acetyltransferase robustness screening.

GlmU acetyltransferase absorbance screening assay controls for robustness screening. Replicate 1 was comprised of microplates 1 to 4, while replicate 2 was comprised of microplates 5 to 8. The uninhibited controls are represented by (●) and the no-enzyme controls, representative of 100% inhibition, are represented by (●). Each of the control data points are a mean of 16 replicates. The Z' factor, determined using Eqn. (24), for each microplate is represented by (●). The solid horizontal lines represent the mean of either the uninhibited control (—), no-enzyme control (—) or the Z' factor (—). The dashed lines represent the +/- standard deviations of each microplate measure.

The GSK screening data quality guidelines were used, so to consider a microplate in the robustness analysis the Z' factor must be more than 0.4. The data in Figure 93 shows that each of the robustness microplates screened in the GlmU acetyltransferase are greater than 0.4 and the mean Z' factor for the entire training set was 0.7 ± 0.1.

The 21 small molecules identified as hits on both test occasions will be re-tested in duplicate using a dilution of each of the compounds.

The second single concentration, small molecule training set screened was the GSK Nuisance set, which is comprised of 1027 compounds that are known to interfere with several assay detection methodologies. These compounds are dispensed in 3, 384 well microplates and are typically screened in duplicate on two separate test occasions to ascertain the reproducibility of the assay. The data presented in Figure 91 is the correlation of the two test occasions of GlmU acetyltransferase absorbance assay.

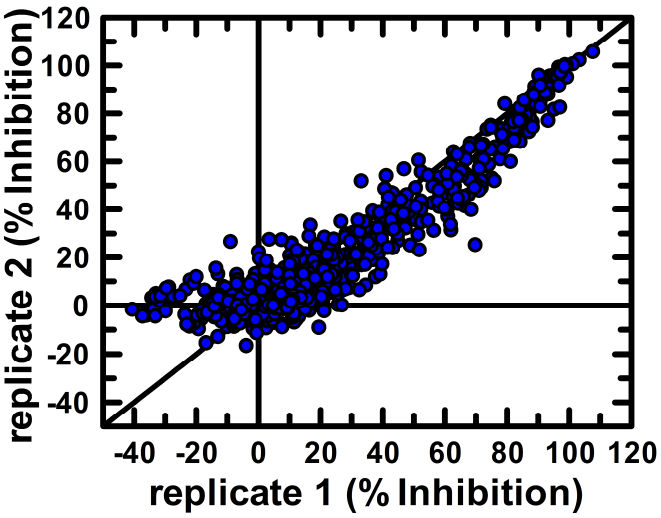


Figure 94. GlmU acetyltransferase assay nuisance set correlation.

GSK nuisance set screened on two separate test occasions and the data normalised to the controls on each of the test microplates using Eqn. (23). The test wells are identified by (●).

The GlmU acetyltransferase activity nuisance screen identified a high number of compounds that are potentially compound-mediated interferences. The statistical threshold is more complex to determine for these data sets as there are many potential hits that skew the distribution of the test compounds in the histograms presented in Figure 95.

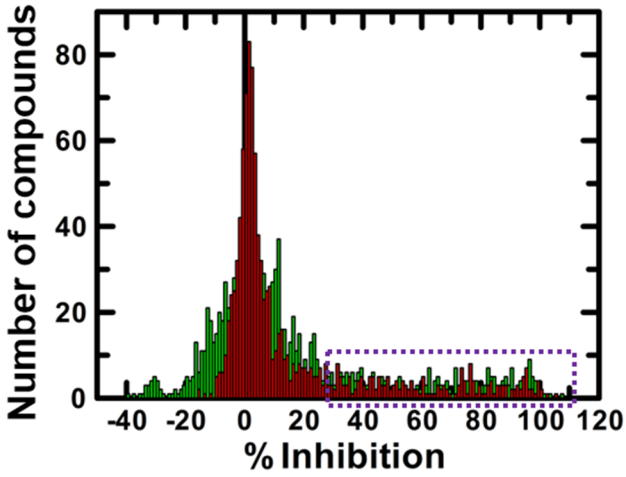


Figure 95. Distribution of % Inhibition for nuisance set test compounds.

Histograms of the distribution of % Inhibition values of the 1000 test compounds of the GSK nuisance set when screened in the GlmU acetyltransferase assay. The histograms in red and green are the % Inhibition data for replicate 1 and 2,

respectively. The dashed purple box highlights the greater number of compounds with % Inhibition greater than 30 % Inhibition.

The data presented in Figure 95 highlights that there are significantly more hits identified when screening the nuisance set, highlighted by the purple dashed box, in comparison to the same data presented for the robustness set (Figure 92). These data confirm that that acetyltransferase DTP absorbance assay is prone to small molecule mediated interference and of the 1027 compounds tested, 227 were identified as hits, which is a hit rate of 22.1%.

The assay performance was assessed by determining the Z' factor from the two control populations of 16 per microplate tested, which were uninhibited (0% Inhibition) and a no-enzyme controls (100% Inhibition). The nuisance set microplate controls and the Z' factor determined from these values are presented in Figure 96.

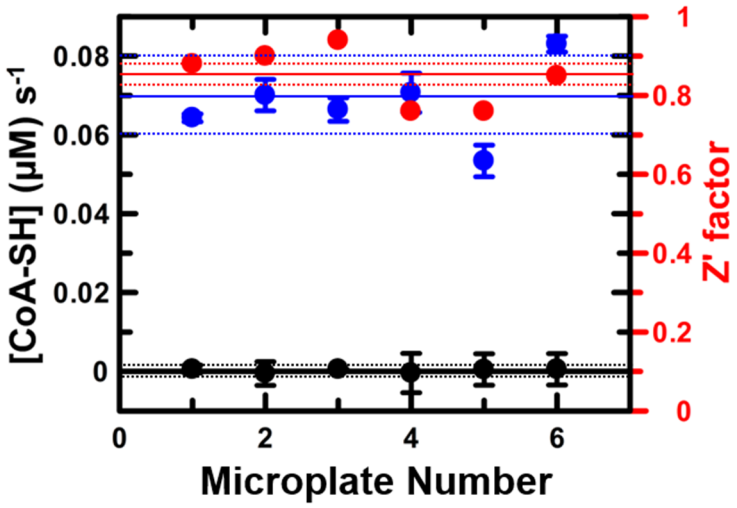


Figure 96. Z' factor and assay controls of GlmU acetyltransferase nuisance set screening.

GlmU acetyltransferase absorbance screening assay controls for nuisance set screening. Replicate 1 was comprised of microplates 1 to 3, while replicate 2 was comprised of microplates 4 to 6. The uninhibited controls are represented by (●) and the no-enzyme controls, representative of 100% inhibition, are represented by (●). Each of the control data points are a mean of 16 replicates. The Z' factor, determined using Eqn. (24), for each microplate is represented by (●). The solid horizontal lines represent the mean of either the uninhibited control (—), no-enzyme control (—) or the Z' factor (—). The dashed lines represent the +/- standard deviations of each microplate measure.

The data presented in Figure 96 shows that each of the nuisance set microplates screened in the GlmU acetyltransferase are greater than 0.4 and the mean Z' factor for the entire training set was 0.85 ± 0.07.

The small molecules that comprise the nuisance set have been annotated with the known assay interference, which provides context to the hits identified by screening this set and subsequently the likely type of interferences that the DTP absorbance assay is susceptible to Table 24.

Table 24. GSK annotations of nuisance set hits

GSK interference mechanism annotation	number of hits
Oxidative stability	40
Redox interference	35
DNA binding interference	31
Free oxygen scavenger	20
Reduction interference	18
Matched Pair Original active sample	13
Matched Pair re-synthesized inactive	13
Protein denaturants	12
Reactive compounds	10
Singlet oxygen quencher	10
Polyphenols	8
Cys reactivity	7
Frequent hit	4

The annotations of the hits identified from screening the GSK nuisance set against the GlmU acetyltransferase activity demonstrates that the DTP absorbance assay is mainly susceptible to compounds that are known to interfere with reduction and oxidation mechanisms. Other compounds identified as frequent hits have mechanisms such as instability, known-protein denaturants and free oxygen

scavengers. The assay methodology, which utilises a mechanism whereby DTP oxidises the thiol group of CoA-SH, is likely to be prone to some of the redox interferences identified from screening this set. The identification of reactive and protein denaturants could also be expected as the absorbance assay utilises low concentrations of GlmU.

5.3 GSK TB set screening

The open source, GSK TB set, detailed in a publication by Ballell, *et al.*[427], now comprising 227 compounds[336], was screened using the GlmU acetyltransferase activity using the DTP absorbance assay. The set was screened in duplicate, on two separate test occasions in the presence of substrate concentrations equivalent to the K_m values (Figure 97).

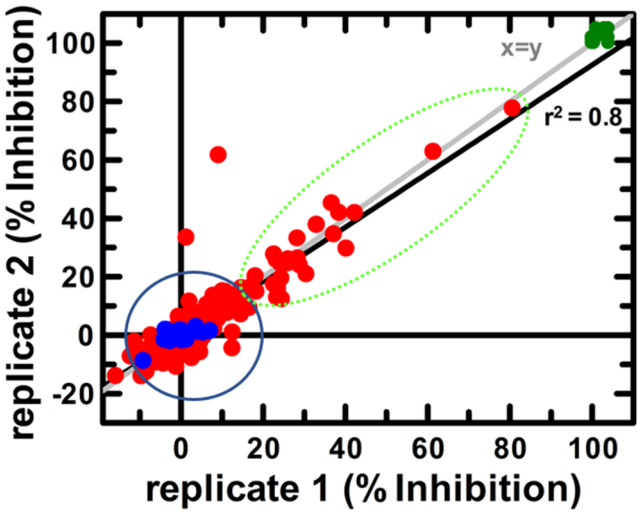


Figure 97. GlmU acetyltransferase assay TB set correlation.

GSK TB set screened on two separate test occasions and the data normalised to the controls on each of the test microplates using Eqn. (23). The uninhibited (●) and no-enzyme (●) control populations are highlighted and the test wells are identified by (●). The blue circle highlights the statistical cut off that determines whether a test compound is identified as a hit or inactive. The green, dashed circle highlights the hit compounds that were selected for dose response testing.

The GlmU acetyltransferase activity TB set screen identified 22 compounds that were statistically classed as hits on both test occasions, which as a percentage of the 227 compounds tested is 9.6%. The distribution of the % Inhibition values of the test compounds is presented in Figure 98.

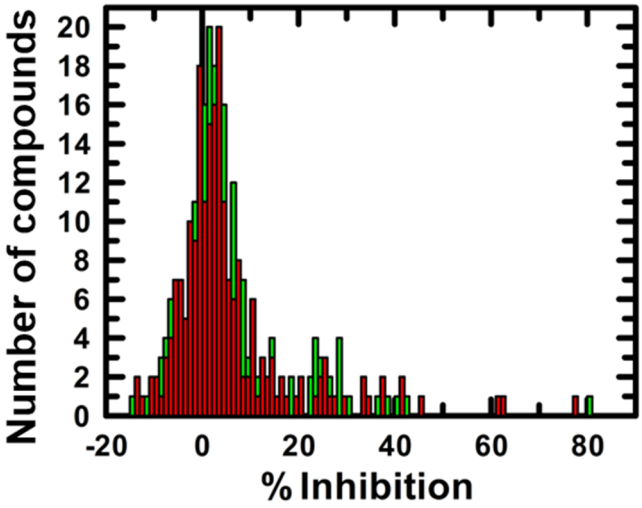


Figure 98. Distribution of % Inhibition for TB set test compounds.

Histograms of the distribution of % Inhibition values of the 211 test compounds of the GSK TB set when screened in the GlmU acetyltransferase assay. The histograms in red and green are the % Inhibition data for replicate 1 and 2, respectively. The data on the y-axis are the count of the values in each histogram category.

The data presented in Figure 98 demonstrates that there are normal distributions for both test occasions of the TB set. The Z' factor for the two TB set microplates tested were 0.87 and 0.62, respectively for replicate 1 and 2.

In parallel to screening of the TB set at a single concentration, the same compound collection was screened as concentration titrations. The TB set concentration titration compounds were prepared as detailed in Section 2.2.22.1 and were dispensed onto 7, 384 well microplates. The concentration titration version of the TB set was screened in duplicate, on two separate test occasions in the presence of substrate concentrations equivalent to the K_m values (Figure 99).

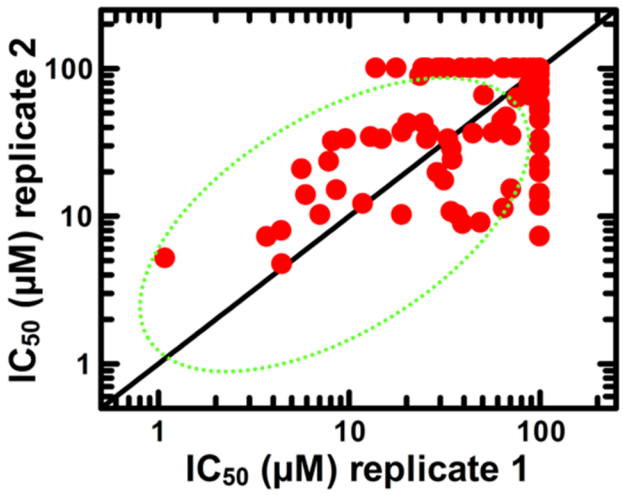


Figure 99. GlmU acetyltransferase assay TB set dose response testing correlation.

GSK TB concentration titration set was screened on two separate test occasions and the data normalised to the controls on each of the test microplates using Eqn. (23). Subsequently, the dose response data for each compound was fitted to Eqn. (8). The green, dashed circle highlights the hit compounds that were selected for further studies.

The data presented in Figure 99 confirms that there are number of compounds in the GSK TB set that appear to inhibit the GlmU acetyltransferase activity, which confirms the single concentration data (Figure 97). There were 39 compounds identified where it was possible to determine an IC_{50} on both test occasions. The assay signal window and Z' factor data for the 14 TB set dose response microplates tested are presented in Figure 100.

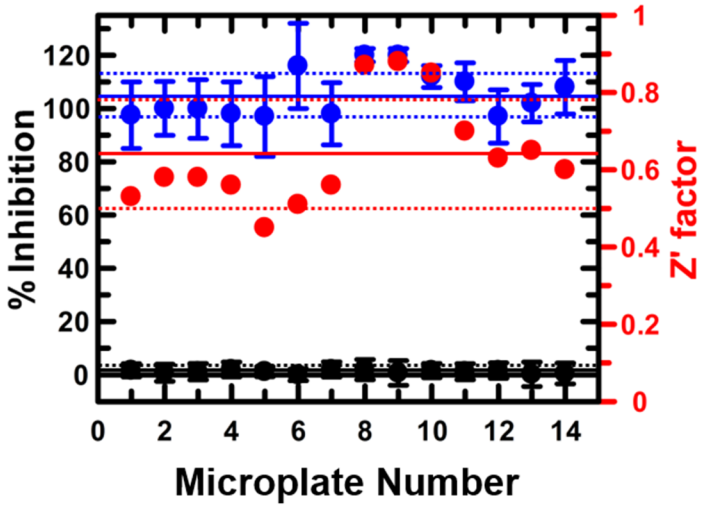


Figure 100. Z' factor and assay controls of GlmU acetyltransferase TB set dose response screening.

GlmU acetyltransferase absorbance screening assay controls for TB set dose response screening. Replicate 1 was comprised of microplates 1 to 7, while replicate 2 was comprised of microplates 8 to 14. The uninhibited controls are represented by (●) and the no-enzyme controls, representative of 100% inhibition, are represented by (●). Each of the control data points are a mean of 16 replicates. The Z' factor, determined using Eqn. (24), for each microplate is represented by (●). The solid horizontal lines represent the mean of either the uninhibited control (—), no-enzyme control (—) or the Z' factor (—). The dashed lines represent the +/- standard deviations of each microplate measure.

The TB set dose response mean Z' factor for the 14, 384 well microplates tested was 0.64 ± 0.14, which is lower than expected and mainly due to the high variability of the uninhibited controls for the microplates of replicate 1. However, the Z' factor for all microplates of the duplicate TB dose response sets tested were greater than the minimum limit of 0.4 and were subsequently involved in the initial correlation analysis (Figure 99).

To confirm both the single concentration and dose response data the two results were correlated against one another to ensure that the hits identified in the single concentration screen were the inhibitors identified in the dose response test occasions (Figure 101).

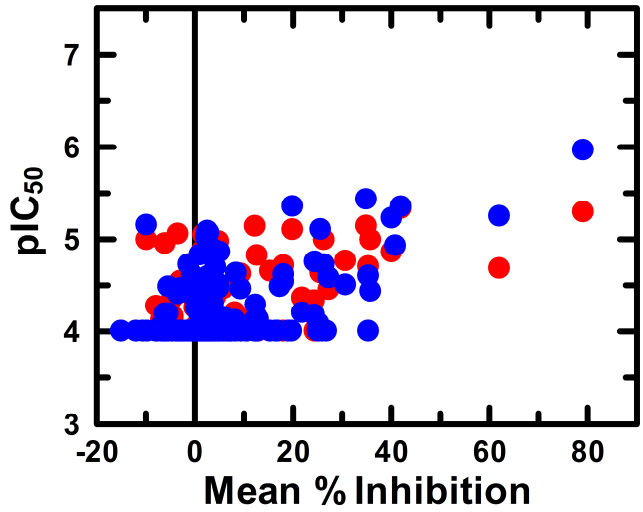


Figure 101. GlmU acetyltransferase assay TB set single concentration test occasions correlated against dose response data.

GSK TB single concentration data, presented as a mean of the two test occasions, correlated against replicate 1 (●) and replicate 2 (●) dose response data, displayed as logIC_{50}, or pIC_{50}.

The data in Figure 101 confirm that most hits identified during single concentration screening were the inhibitors found during dose response testing with IC_{50} values lower than 100 µM. The IC_{50} and pIC_{50} data for the confirmed TB set hits are detailed in Table 25.

Table 25. GlmU acetyltransferase TB set hits [a].

GSK compound number [b]	Mean IC_{50} (µM)	Mean pIC_{50}
GSK810037A	3.1	5.6
GSK1611550A	4.6	5.3
SB-435634	5.5	5.3
GSK1826089A	6.2	5.2
GSK735826A	8.6	5.1
GSK920684A	9.9	5.0
AH24761A	13.2	5.0
GSK1826825A	11.8	4.9

GSK426032A	12.0	4.9
GSK690382A	15.6	4.9
GSK1385423A	14.6	4.9
GSK1905227A	20.1	4.8
GSK1829674A	21.4	4.7
GSK1329151A	24.2	4.7
GSK1180781A	22.6	4.7
GSK735816B	23.6	4.7
GR223839X	23.6	4.7
GSK920703A	29.0	4.7
CCI7967	24.0	4.7

[a] Data are a mean of two replicates. [b] GSK registration numbers for compounds included in the GSK TB set publication by Ballel, *et al.*[427]

The compounds listed in Table 25, as well as the hits from the robustness set screening (Figure 91) were re-prepared from alternative stocks of compounds and re-tested as concentration titrations in the DTP absorbance assay (Figure 102).

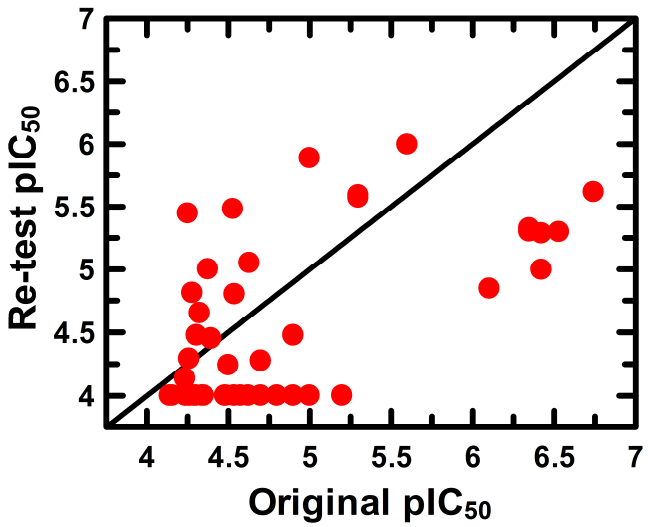


Figure 102. GlmU acetyltransferase assay TB and robustness set hit dose response confirmation.

Concentration response test data of re-prepared hit compound concentration titrations, presented as a scatter plot. The original pIC_{50} test values are presented on the x-axis and the pIC_{50} values determined from the re-prepared concentration titrations are on the y-axis. Data normalised to microplate controls Eqn. (23) and then fitted to Eqn. (8). The x=y trend is represented by (—).

Several potential hits identified from screening the robustness and TB sets did not re-confirm when re-tested from new preparations of test compounds, which was not unexpected, as preparations of test compounds can be unstable. The more potent hits identified from screening did re-test, albeit with different pIC_{50} values from the original test value.

The inhibitors that re-confirmed in the dose response testing experiment, utilising new preparations of compounds, were assessed further. The most promising compound, GSK920684A, was further investigated along with three related compounds. These compounds were identified and then tested, as concentration titrations in the GlmU acetyltransferase activity DTP absorbance assay, alongside a further retest of the original hit (Figure 103).

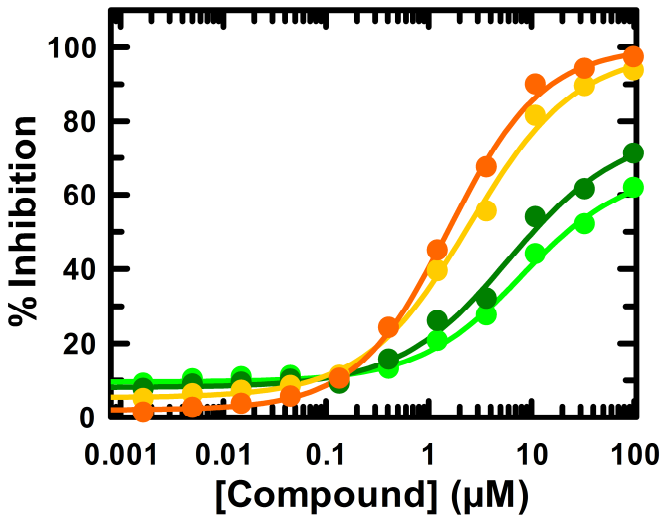


Figure 103. GlmU acetyltransferase assay dose response testing of TB set hit and close analogues.

TB set hit, GSK920684A (●), re-tested in the DTP absorbance assay, along with three close analogues; GSK024A (●), GSK025A (●), and GSK026A (●). These data were normalised to microplate controls using Eqn. (23) and subsequently fitted to Eqn. (8).

The dose response inhibition of TB set hit, GSK920684A, re-tested and an IC_{50} of 2.5 ± 0.3 µM was determined. The three close analogues of GSK920684A tested concurrently, were related to the original TB set hit by Tanimoto similarities of between 0.93 to 0.96. The only analogue to inhibit 100 % of GlmU acetyltransferase activity was GSK026A, which was determined to have an IC_{50} value of 1.6 ± 0.1 µM. The other analogues did not inhibit 100 % of GlmU acetyltransferase activity at the concentrations utilised for this study, however, fitting the data to Eqn. (8) allowed determination of apparent IC_{50} values of 6.1 ± 1.7 µM and 9.1 ± 1.9 µM, for GSK024A and GSK025A, respectively.

The compounds tested in the GlmU acetyltransferase activity assay were subsequently tested in a DSF thermal stability assay to ascertain whether these potential inhibitors stabilised GlmU T_m (Table 26).

Table 26. GlmU DSF studies in the presence of inhibitors identified from TB set screening [a].

GSK compound number	Maximum ΔT_m (°C) [b, c]	K_d (µM) [b, c]
GSK920684A	1.3 ± 0.1	2.2 ± 0.2
GSK024A	0.2 ± 0.08	N.D.
GSK025A	0.6 ± 0.12	N.D.
GSK026A	0.9 ± 0.1	35.3 ± 11

[a] Experiments carried out using 1.5 µM GlmU and 10 x SYPRO Orange. Compounds incubated with GlmU for 15 mins at 25 °C, prior to addition of SYPRO Orange and application of the thermal gradient. Compounds tested from 200 µM and were titrated by 1.5-fold. [b] Data are a mean of four replicates. [c] Data were fitted to Eqn. (19) first to determine T_m and ΔT_m before fitting to Eqn. (20) to determine K_d.

The data presented in Table 26 demonstrates that GlmU T_m was stabilised from 0.9 °C or higher by GSK920684A and GSK026A. The dose response data for these compounds were fitted to Eqn. (20) and subsequently K_d values were determined for both GSK920684A and GSK026A. The GlmU T_m was only modestly stabilised by either 0.2 or 0.6 °C in the presence of either GSK024A or GSK025A, respectively.

This study was controlled by 16 replicates of non-liganded GlmU, which, when fitted to Eqn. (19), yielded a T_m of 51.7 ± 0.1 °C, in addition to two titrations of Ac-CoA from 3 mM. The GlmU T_m in the presence of Ac-CoA was stabilised by a maximum T_m of 4.2 ± 0.1 °C and this dose response data when fitted to Eqn. (20) yielded a K_d value of 160 ± 17 µM. These control results confirm that the GlmU DSF data were as expected when compared to previous studies in experiment Table 9 and Figure 57.

To further confirm the validity of these GlmU inhibitors an alternative assay was required to monitor acetyltransferase activity, which would ideally not utilise either a thiol detection or absorbance assay methodology, as to discount the possibility of interference mechanisms. RF-MS was identified as a potential method to directly monitor the simultaneous reduction in concentration of the substrate, GlcN-1P, in addition to the increase in amounts of the second product, GlcNAc-1P. Initially, GlcN-1P and GlcNAc-1P were titrated in assay buffer adjusted to pH 7.5, prior to addition of 1% formic acid and detection of substrate and product using RF-MS (Figure 104).

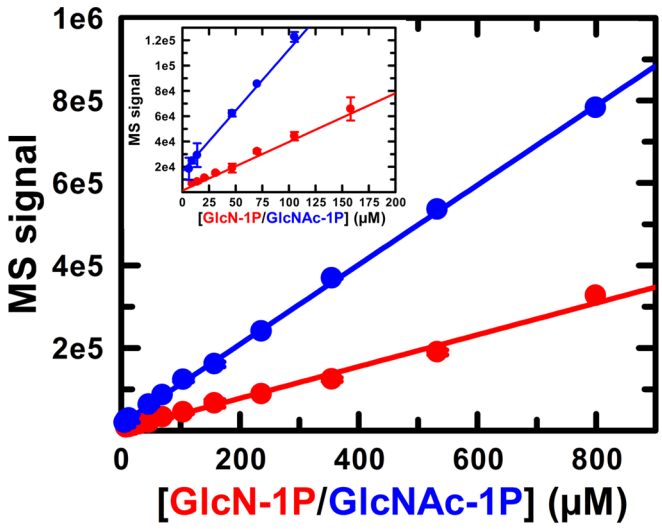


Figure 104. RF-MS detection of GlcN-1P and GlcNAc-1P titrations.

GlcN-1P (●) or GlcNAc-1P (●) were titrated in assay buffer at pH 7.5 and subsequently detected by RF-MS, following additions of 1% Formic acid. The y-axis data are the smoothed, extracted ion-chromatogram (XIC) peak areas for each compound. Each of these data have been fitted to a linear regression; GlcN-1P (—, correlation coefficient = 0.99) or GlcNAc-1P (—, correlation coefficient = 0.99). The inset plot presents the titration data of GlcN-1P and GlcNAc-1P below a concentration of 200 µM.

The data presented in Figure 104 demonstrate that it is possible to use RF-MS to monitor a linear concentration response of GlcN-1P and GlcNAc-1P from 2 µM up to

800 µM. The minimum level of detection of GlcN-1P and GlcNAc-1P is equivalent to 1% substrate turnover when using 200 µM substrate. Subsequently, after validating the use of RF-MS to simultaneously detect the relative concentrations of GlcN-1P and GlcNAc-1P from the same microplate well, the next experiment was to ascertain the linearity GlmU acetyltransferase activity. This was achieved by titrating enzyme and initiating acetyl transfer by addition of 200 µM GlcN-1P and Ac-CoA before stopping the reaction by addition of 0.5% Formic acid at several discrete time points and then detecting concentrations of substrate and product using RF-MS (Figure 105).

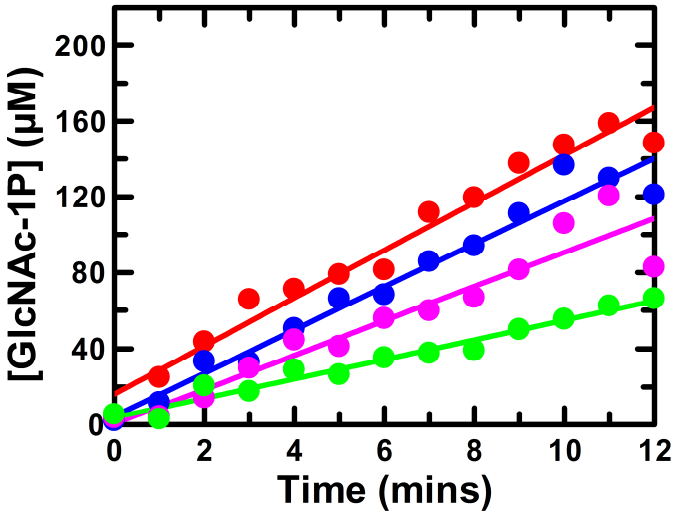


Figure 105. RF-MS detection of GlcNAc-1P formed by GlmU acetyltransferase.

These experiments were carried out with the help of Michelle Pemberton, Screening, Profiling and Mechanistic Biology, GSK. GlmU was titrated in the presence of 200 µM of both Ac-CoA and GlcN-1P in assay buffer at pH 7.5. Accumulation of the product was monitored by RF-MS and normalised to a GlcNAc-1P standard curve, on the same microplate. The concentrations of GlmU used were; 5 nM (●), 2.5 nM (●), 1.25 nM (●) and 0.613 nM (●).

GlmU acetyltransferase activity was linear for all concentrations of enzyme tested in Figure 105, and a maximum of 160 µM of GlcNAc-1P was formed by 5 nM GlmU after 12 mins, which is equivalent to 80% substrate turnover. These data are in broad agreement with the substrate turnover of 5 nM GlmU acetyltransferase determined using the DTP absorbance assay, which approximated that 50% CoA-SH was formed after 8.3 mins (Figure 89). Prior to finalising the GlmU acetyltransferase RF-MS assay conditions, the steady state parameters were determined for the

acetyltransferase substrates by titrating either GlcN-1P or Ac-CoA in the presence of a fixed, saturating concentration of the second substrate and 5 nM GlmU (Figure 106).

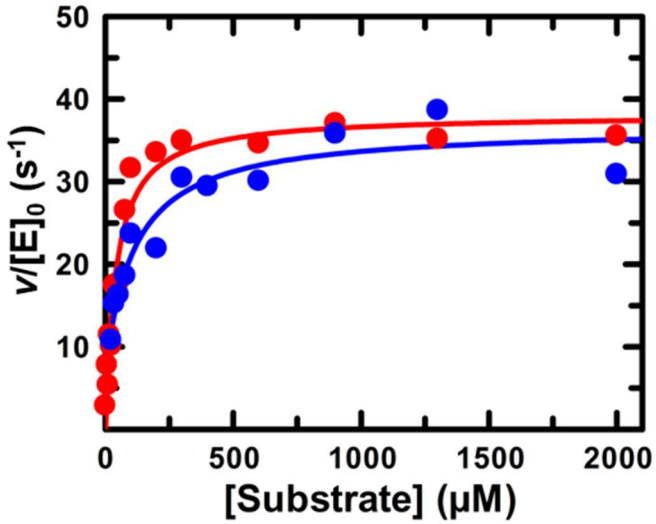


Figure 106. GlmU acetyltransferase RF-MS steady-state kinetics.

These experiments were carried out with the help of Michelle Pemberton, Screening, Profiling and Mechanistic Biology, GSK. Initial velocity data for GlmU acetyltransferase reaction using titrations of either Ac-CoA (●) or GlcN-1P (●) in the presence of saturating concentrations of the second substrate. These discontinuous data were fitted to a linear regression model and then normalised using a GlcNAc-1P standard curve, on the same microplate. Symbols represent experimental data, and solid lines the fit of the data to Eqn. (1), for either Ac-CoA (—) or GlcN-1P (—). These data are a mean of two discrete replicates.

The acetyltransferase activity of GlmU was monitored in the presence of a titration of Ac-CoA, at a saturating concentration of GlcN-1P enabled determination of a $K_{m,\ Ac\text{-}CoA}$ of 112.6 ± 32 μM and $k_{cat,\ Ac\text{-}CoA}$ = 29.6 ± 3.9 s^{-1}. The reverse experiment, using a titration of GlcN-1P in the presence of a high concentration of Ac-CoA was also fit to Eqn. (1), which gave a $K_{m,\ GlcN\text{-}1P}$ of 40.9 ± 5.8 μM and $k_{cat,\ GlcN\text{-}1P}$ = 38.1 ± 1.3 s^{-1}. The GlmU acetyltransferase Michaelis-Menten parameters using the RF-MS assay, which monitored accumulation of GlcNAc-1P, were broadly in agreement with the steady state values determined using the DTP absorbance assay. Inhibition of GlmU acetyl transfer monitored using the RF-MS assay was first validated utilising CoA-SH to confirm product inhibition, which was presented previously in Figure 58.

The GlmU acetyltransferase activity RF-MS assay was subsequently used to confirm the inhibition of the top TB set hit and the three analogues, initially investigated using the DTP absorbance assay in Figure 103. The inhibitors were titrated and then

incubated with GlmU, prior to initiation of acetyl transfer with 200 µM Ac-CoA and GlcN-1P. The acetyltransferase reaction was stopped using 0.5% formic acid and quantification of the concentration of GlcNAc-1P using RF-MS (Table 27).

Table 27. GlmU RF-MS inhibition studies using inhibitors identified from TB set screening [a].

GSK compound number	RF-MS IC_{50} (µM) [b, c]	DTP absorbance IC_{50} (µM) [c, d]
GSK920684A	49.9 ± 4.6	42.7 ± 3.5
GSK024A	37.7 ± 4.1	106.8 ± 30.1
GSK025A	250.5 ± 96.3	430.3 ± 8.3
GSK026A	89.1 ± 15.6	80.9 ± 10.3

[a] Experiments carried out using 2.5 nM GlmU, 200 µM Ac-CoA and GlcN-1P. Compounds incubated with GlmU for 15 mins at 25 °C, prior to addition of substrates to initiate the reaction. An addition of 0.5% formic acid stopped the reaction. Compounds tested from 100 µM and were titrated by 2-fold. [b] Data are a mean of five replicates. [c] Data were fitted to Eqn. (23) first to determine % Inhibition, before fitting to Eqn. (8) to determine IC_{50}. [d] Comparator data, generated simultaneously using the DTP absorbance assay. These data are a mean of two replicates.

The data in Table 27 confirms that the inhibitor, GSK920684A, identified by screening the TB set and the three analogues all inhibit GlmU acetyltransferase activity in an assay independent manner. There is a significant discrepancy, 25 to 50-fold, between the IC_{50} values determined using RF-MS when compared to the values determined, initially, using the DTP absorbance assay (Figure 103). GSK920684A and analogues were retested simultaneously using the DTP absorbance assay and the IC_{50} values determined were much more comparable to the results of testing these compounds in the RF-MS assay (Table 27). This could potentially be due to the use of re-purified compounds, which were prepared from solid immediately before carrying out the RF-MS and DTP absorbance assays. There were some general trends between all the datasets, primarily that the magnitude of inhibition of GSK920684A and GSK026A was greater than the values achieved with GSK024A and GSK025A. It should be noted that both the starting

concentration and dilution series of all the test compounds was inappropriate and should have been both higher and lower, in terms of dilution factor. These data in combination with the DSF direct binding data confirm that GSK920684A, and analogues, bind to GlmU acetyltransferase domain and inhibit the enzyme activity.

To further explore this compound series, the GSK compound collection was searched to identify related compounds that could be tested as inhibitors against the GlmU acetyltransferase activity. The compound search identified 81 small molecules (with a Tanimoto similarity equal to 0.8 or greater) that were tested in a series of dose response experiments. These studies utilised titrations of analogues from 100 µM and were subsequently tested against the GlmU acetyltransferase activity, using the DTP absorbance assay, in the presence of 200 µM Ac-CoA and GlcN-1P (Figure 107).

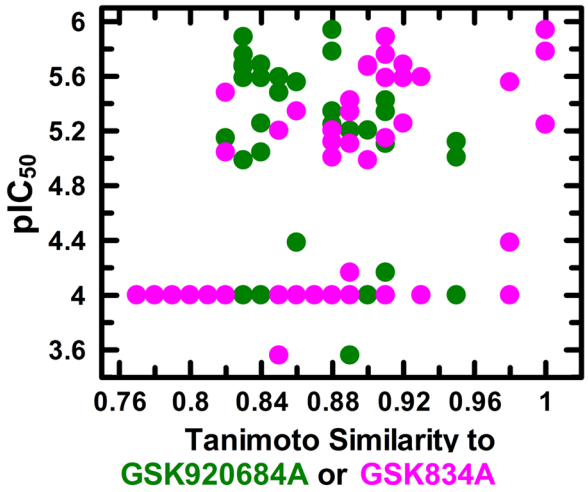


Figure 107. Dose response testing of GSK920684A analogues against GlmU acetyltransferase activity.

Concentration titrations of GSK920684A analogues tested in the GlmU acetyltransferase absorbance assay in assay buffer adjusted to pH 7.5. The dose response data were fitted to Eqn. (8), and the pIC_{50} values (●) were then correlated against the Tanimoto similarity of the analogues to GSK920684A. Alternatively, the pIC_{50} values (●) were also correlated against the Tanimoto similarity of the analogues to GSK834A, which is the most potent compound.

The dose response data of GSK920684A analogues, presented in Figure 107 as the fitted pIC_{50} value demonstrate that there is no correlation between similarity to the original TB set hit and the tested pIC_{50} (r^2 value of -0.4). An alternative correlation is presented in Figure 107 between the pIC_{50} and the Tanimoto similarity of test compounds to the most potent inhibitor, GSK834A, which yielded an r^2 value of 0.5.

The correlation coefficient of pIC_{50} with the Tanimoto similarity to GSK834A suggests that there is a compound structure activity relationship (SAR) that could be explored to optimise the potency of these inhibitors. In this study, 81 compounds were tested and 25 had pIC_{50} values of 4.95 or higher, which further confirms this series of small molecules are promising inhibitors of the GlmU acetyltransferase activity.

To confirm the binding of GSK920684A analogues, to GlmU, MST direct interaction studies were carried out. MST is a biophysical technique that utilises fluorophore labelled protein (referred to from hereon in as fluoro-protein) to monitor changes in the migration of fluoro-protein upon application of a limited temperature gradient to the sample. The migration of the fluoro-protein is modulated by the concentration and affinity of the test compound, allowing determination of a K_d for the binding interaction by first fitting the data to Eqn. (25) and then to Eqn. (8). Prior to testing the GSK920684A analogues, GlmU was labelled with the NanoTemper fluorophore, NT647, following the protocol detailed in Section 2.2.27. To initially validate MST as an appropriate methodology to monitor direct binding to GlmU, Ac-CoA was titrated from 10 mM in the presence of 46 nM of fluoro-labelled protein, prior to aspiration of the samples into capillaries and then measurement of a MST signal (Figure 108).

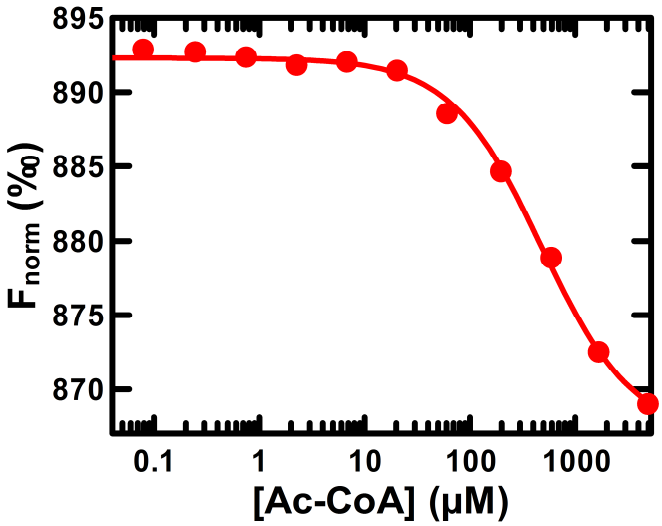


Figure 108. MST studies investigating binding of Ac-CoA to GlmU.

These experiments were carried out with the help of Sabrina Bedard, Protein, Cellular and Structural Sciences, GSK. Titrations of Ac-CoA in the presence of 46 nM GlmU. The data are a mean of 3 discrete replicates. The data for each condition investigated were first fitted to Eqn. (25), and then Eqn. (26) to determine the K_d.

The Ac-CoA binding studies had a concentration response, which when fitted to the Eqn. (26) yielded a K_d value of 480 ± 40 µM. The GlmU $K_{d,\ Ac\text{-}CoA}$ determined using MST agrees with the K_d value determined using DSF (Figure 57), which gives confidence in both the $K_{d,\ Ac\text{-}CoA}$ and MST for characterising binding to GlmU.

The GlmU MST binding assay was then utilised to investigate five compounds that are analogues of the initial TB set hit, GSK920684A. The compounds were titrated from 500 µM in the presence of 50 nM fluoro-labelled GlmU in a microplate, prior to aspiration into test capillaries. The experimental protocol utilised differed from the method used for Ac-CoA binding study, primarily due to a change in fluorescence observed from the test capillaries, prior to initiating the laser heating and measurement of the MST signal. This is a commonly observed phenomenon and can be due to modulation of the fluorescence of the fluorophore labelled protein, in the presence of compound[434,435]. The change in initial observed fluorescence was utilised to calculate the fraction of test compound bound to fluoro-labelled GlmU, which subsequently can be used to determine a EC_{50} value (Table 28).

Table 28. GlmU MST binding studies using GSK920684A analogues [a].

GSK compound number	EC_{50} (µM) [c]	$IC_{50,}$ (µM) [d]	Similarity to GSK920684A (Tanimoto score)
GSK102A	13.2 ± 1.1	21.3 ± 26.1	0.86
GSK105A [b]	10.7 ± 1.5	2.4 ± 0.4	0.84
GSK104A [b]	10.7 ± 2.8	2.2 ± 0.04	0.83
GSK103A	7.9 ± 1.0	1.6 ± 0.3	0.83
GSK984A	151 ± 46	11.8 ± 5.6	0.91

[a] Experiments carried out using 50 nM NT647 labelled-GlmU. Compounds incubated with GlmU for 15 mins at 25 °C, prior to aspiration into test capillaries. [b] Data are a mean of two replicates. [c] Data were fitted to Eqn. (25) first to determine a fraction bound, before fitting to Eqn. (27) to determine EC_{50}. [d] GlmU acetyltransferase utilising the DTP absorbance assay. Compounds were incubated with 5 nM GlmU for 15 mins at 25 °C, prior to initiation of acetyl transfer by the addition of 200 µM Ac-CoA and GlcN-1P. Data were normalised to microplate controls using Eqn. (25) and then fitted to Eqn. (27).

The experiments in Table 28 were carried out with the help of Sabrina Bedard, Protein, Cellular and Structural Sciences, GSK. The GSK920684A analogues assessed in the MST direct binding assay were all confirmed as GlmU interacting compounds. The MST binding experiments were carried out in the absence of both acetyltransferase substrates, which indicates that the test compounds could all bind to apo-GlmU. As the binding of the test compounds was determined by monitoring a change in fluorescence from the fluoro-labelled GlmU, additional control experiments were required. The control assessments were carried out to ascertain that the observed compound induced fluorescence quench was not due to aggregation of the fluoro-labelled GlmU. These control experiments known as an "SD test" and involve preparation of mixtures containing the test compound concentrations and 50 nM NT647-GlmU in a PCR microplate, prior to centrifugation for 10 min at 14000 rpm. This centrifugation step removed any protein precipitate before addition of two-fold concentrate SD solution, which contained 4% SDS and 40 mM DTT. The mixtures were subsequently transferred to centrifuge tubes and heated at 90 °C for 15 mins, before an additional centrifugation step at 14000 rpm for 1 min. The mixtures were then aspirated into capillaries and fluorescence detected using the Monolith NT.115 instrument. There was no change in fluorescence for any of the compound concentrations tested (for all five GSK920684A analogues tested), which is indicative of compound-induced fluorescence quenching of NT647-GlmU and mitigates any concerns regarding compound-induced protein denaturation. If compound-induced GlmU denaturation was present, then the fluorescence change monitored from the SD test samples should be the same as the values utilised to calculate the EC_{50} values in Table 28. The SD test confirms that the observed quench of fluorescence from NT647-GlmU is compound-mediated and suggests that this occurs upon compound binding.

The data in Table 28 demonstrate that for all compounds, apart from GSK984A, have EC_{50} values in the range of 7.9 to 13.2 μM. There is a consistent 5-fold difference between the MST EC_{50} and absorbance assay IC_{50} values for all compounds, apart from GSK984A, which conversely is the most similar compound to GSK920684A. The apparent IC_{50} values are more potent than the EC_{50} values, which could be explained by the presence of concentrations of the acetyltransferase substrates that are equivalent to the measured K_m values in the acetyltransferase activity assay. To

provide a positive control for the compound MST studies, the Ac-CoA binding experiment was repeated concurrently, which yielded a K_d value of 330 ± 70 µM.

In addition to the 227 compound, GSK TB set, an additional collection of 3000 *Mycobacterium tuberculosis* targeted small molecules, known as the TB box. These small molecules were identified from the same HTS as the compounds comprising the TB set but unlike the smaller collection, the TB box was kept as a confidential GSK resource. The aim of screening the TB box collection of small molecules is to identify further compounds or series of inhibitors that could be utilised as either tool molecules to better understand the role of GlmU in cell wall biosynthesis or provide starting points for future inhibitor optimisation.

The TB box set of compounds were screened at a single concentration of 10 µM in the GlmU acetyltransferase DTP absorbance assay, utilising assay buffer adjusted to pH 7.5 and 200 µM of Ac-CoA and GlcN-1P (Figure 109).

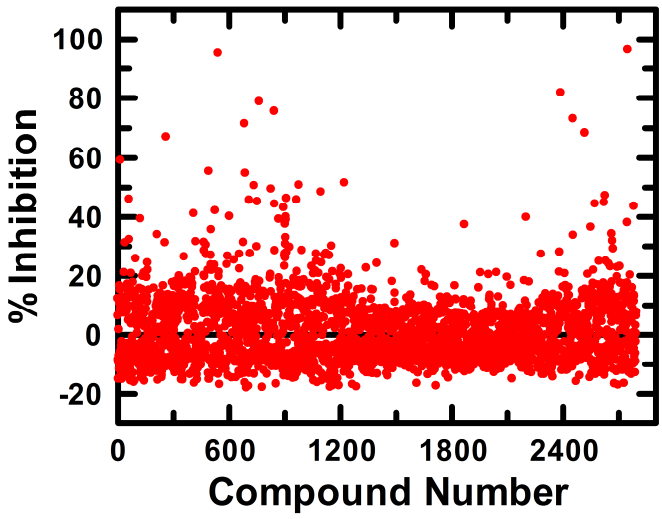


Figure 109. GlmU acetyltransferase screening of TB box small molecule set.

GlmU acetyltransferase activity screening of the GSK TB box small molecule collection. The initial rate data were normalised to the microplate control populations using Eqn. (23). The test compound number is on the x-axis and the % Inhibition, determined from the control populations on each test microplate, is on the y axis.

The data presented in Figure 109 demonstrates that there were several potential GlmU acetyltransferase hit compounds identified from the initial single concentration screen of the TB box set. The statistical hit threshold, determined from the mean and standard deviation of the non-inhibitory test compounds, was 26.5 % Inhibition, which yielded a hit rate of 2.6 %, or 77 potential inhibitors. The assay signal window

and Z' factor data for the 7 TB box set microplates tested are presented in Figure 110.

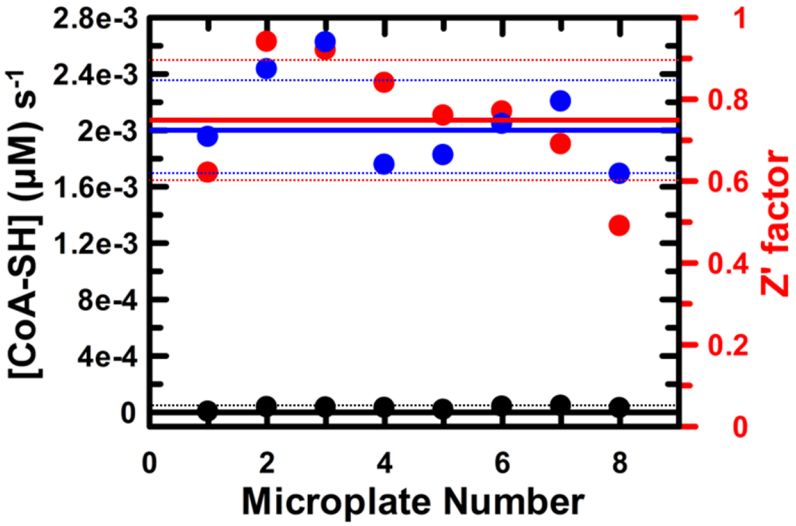


Figure 110. Z' factor and assay controls of GlmU acetyltransferase TB box set single concentration screening.

GlmU acetyltransferase absorbance screening assay controls for TB box set single concentration screening. The uninhibited controls are represented by (●) and the no-enzyme controls, representative of 100% inhibition, are represented by (●). Each of the control data points are a mean of 16 replicates. The Z' factor, determined using Eqn. (24), for each microplate is represented by (●). The solid horizontal lines represent the mean of either the uninhibited control (—), no-enzyme control (—) or the Z' factor (—). The dashed lines represent the +/- standard deviations of each microplate measure.

The TB box set single concentration mean Z' factor for the 7, 384-well microplates tested was 0.75 ± 0.15, which was as expected and greater than the minimum limit of 0.4 and all microplate data were subsequently involved in the hit identification analysis (Figure 109).

The 77 potential inhibitors identified from screening the GlmU acetyltransferase activity against the TB box single concentration set were re-tested as concentration titrations from 10 mM stocks of compounds (Figure 111).

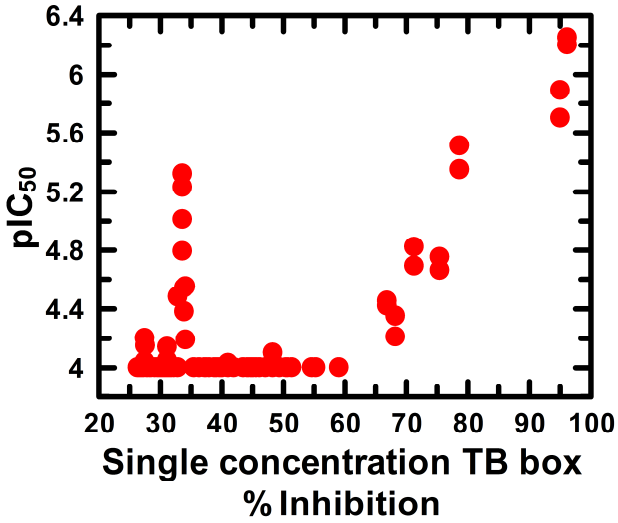


Figure 111. Dose response testing of GlmU TB box hit compounds.

TB box hit compounds, re-prepared as concentration titrations from 10 mM stocks and then tested against the GlmU acetyltransferase activity DTP absorbance assay. The absorbance assay was carried out in assay buffer adjusted to pH 7.5, utilising 5 nM GlmU, 400 µM DTP, 200 µM Ac-CoA and GlcN-1P. The dose response data was normalised to the on-microplate controls and then fitted to Eqn. (8). The data is presented as a scatter plot with the TB box single concentration % inhibition on the x-axis and the pIC_{50} value on the y-axis.

The general trend when testing the 77 TB box hits in the dose response experiment was that compounds with a single concentration response in excess of 60 % inhibition were confirmed as inhibitors. There were 14 confirmed inhibitors with % inhibition values greater than 60 and pIC_{50} values more than 4.2. The two most potent GlmU acetyltransferase inhibitors identified from small molecule screening were identified from the TB box set, with pIC_{50} values of 6.20 and 6.25 (IC_{50} values of 562 and 631 nM). Most of the hits that had % inhibition responses below 60, were not confirmed as inhibitors, however, there were a number of compounds between 27 and 48 % inhibition that when fitted to Eqn. (8) yielded a pIC_{50} in excess of 4. The range of pIC_{50} values for the compounds with single concentration responses below 60 % inhibition was from 4.1 to 5.3.

The data in Figure 111 confirms that there are yet more inhibitors that can be used to further understand the role of GlmU acetyltransferase activity, in addition to, offering several potential starting points for future SAR chemistry. The confirmed inhibitors identified from the TB box small molecules are not related chemically to the TB set hit, GSK920684A (no Tanimoto similarity greater than 0.7).

5.4 GSK validation set screening

At GSK, prior to embarking on a full HTS screen of more than 2 million compounds, a smaller set of 10000 compounds is screened as a final assessment of the suitability of the chosen assay. The 10000-small molecule set is known as the validation set and this collection of compounds represents a significant percentage of the diversity present in the GSK HTS collection. GlmU acetyltransferase activity was screened against the GSK validation set using the DTP absorbance assay to ascertain whether it was possible to identify any novel, hit compounds. In addition, using the DTP absorbance assay to screen the validation set will also allow an initial assessment of the suitability of this approach for future single concentration screens of the GlmU acetyltransferase activity. The validation set was screened utilising the absorbance assay in assay buffer adjusted to pH 7.5, using 5 nM enzyme, 400 µM DTP, 200 µM of Ac-CoA and GlcN-1P (Figure 112).

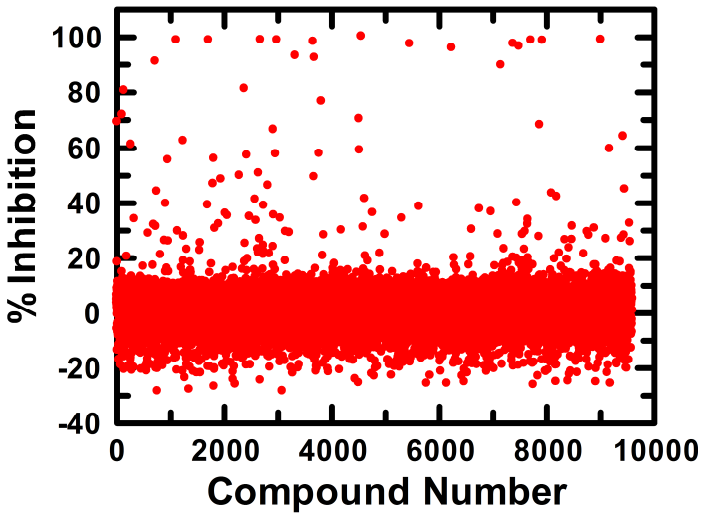


Figure 112. GlmU acetyltransferase screening of the GSK validation small molecule set.

GlmU acetyltransferase activity screening of the GSK validation small molecule collection. The initial rate data were normalised to the microplate control populations using Eqn. (23). The test compound number is on the x-axis and the % Inhibition, determined from the control populations on each test microplate, is on the y axis.

The data presented in Figure 112 demonstrates that there were several potential GlmU acetyltransferase hit compounds identified from the initial single concentration

screen of the GSK validation set. The statistical hit threshold was determined using the normal distribution of the non-inhibitory test compounds (Figure 113).

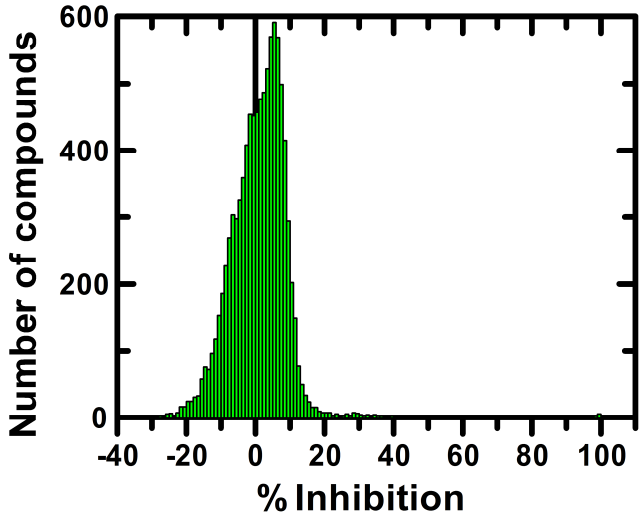


Figure 113. Distribution of % Inhibition for validation set test compounds.

Histograms of the distribution of % Inhibition values of the 10000 test compounds of the GSK validation set when screened in the GlmU acetyltransferase assay. The data on the y-axis are the count of the values in each histogram category.

The mean response of the validation test compounds after normalisation to the on-microplate assay controls was 1.48 ± 9.5% inhibition, which yielded a hit classification threshold of 17% inhibition (after robust weighting of hit compounds). Application of the statistically derived hit threshold led to classification of 168 compounds as potential inhibitors, which is a hit rate of 1.68%. The assay signal window and Z' factor data for the 28 validation set microplates tested are presented in Figure 114.

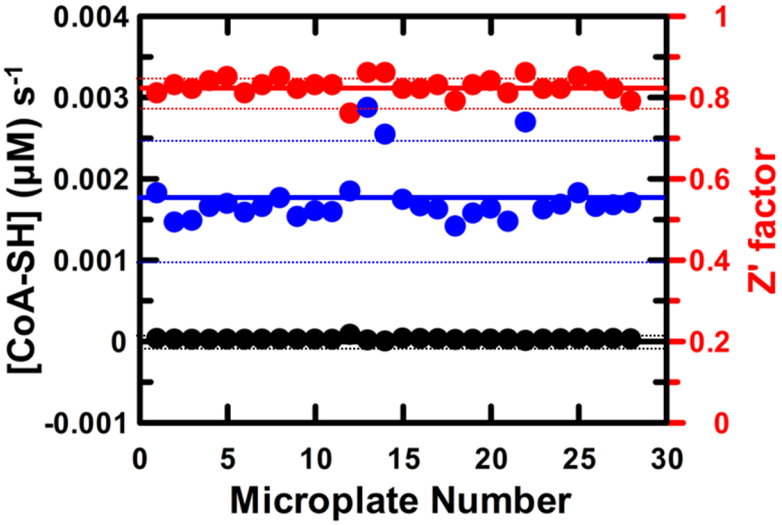


Figure 114. Z' factor and assay controls of GlmU acetyltransferase validation set single concentration screening.

GlmU acetyltransferase absorbance screening assay controls for validation set single concentration screening. The uninhibited controls are represented by (●) and the no-enzyme controls, representative of 100% inhibition, are represented by (●). Each of the control data points are a mean of 16 replicates. The Z' factor, determined using Eqn. (24), for each microplate is represented by (●). The solid horizontal lines represent the mean of either the uninhibited control (—), no-enzyme control (—) or the Z' factor (—). The dashed lines represent the +/- standard deviations of each microplate measure.

The validation set single concentration mean Z' factor for the 28, 384-well microplates tested was 0.83 ± 0.02, which was as expected and greater than the minimum limit of 0.4 and all microplate data were subsequently involved in the hit identification analysis (Figure 112 and Figure 113).

The 168 potential inhibitors identified from screening the GlmU acetyltransferase activity against the validation single concentration set will be re-tested later to determine IC_{50} values for these hits.

5.5 Literature-based inhibitor identification and further dose response testing

As previously discussed in Section 1.4.5, there have been several efforts to identify novel chemical entities that inhibit both *M. tuberculosis* GlmU, as well as several orthologues. To investigate whether the novel chemical entities are inhibitors of *M. tuberculosis* GlmU, a sub-section of related small molecules were selected from the

GSK compound collection. The testing of literature reported inhibitors focussed on the compounds reported by two studies; a GSK encoded library technology (ELT) screen of multiple *M. tuberculosis* proteins that identified a small number of GlmU-interacting molecules and an Astra Zeneca report of a series of compounds identified from a HTS conducted against the *E. coli* GlmU orthologues acetyltransferase activity.

The GlmU interacting small molecules that were identified from the GSK ELT screen were assessed along with similar compounds, were selected to screen against the acetyltransferase activity. To assess the ELT-related small molecules, 148 compounds were titrated from 10 mM stocks and incubated with 5 nM GlmU, prior to initiation of acetyl transfer by the addition of 200 µM Ac-CoA and GlcN-1P, followed by detection of CoA-SH using the DTP absorbance assay (Figure 115).

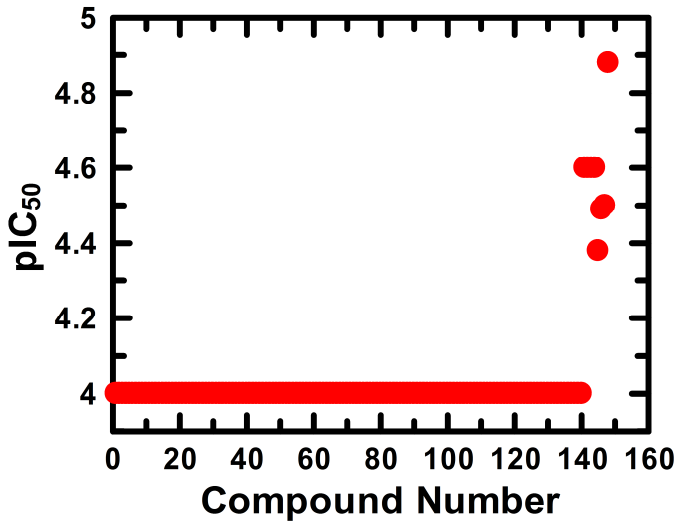


Figure 115. GlmU acetyltransferase screening of compounds related to the proposed ELT pharmacophore.

GlmU acetyltransferase screening of compounds related to the small molecules identified by ELT. The initial rate data were normalised to the microplate control populations using Eqn. (23), before fitting these data to Eqn. (8) to determine the compound pIC_{50} values. The pIC_{50} values for each of the test compounds are represented by (●). The compound number on the x-axis is arbitrary and utilised to demonstrate the spread of pIC_{50} values.

There were 8 compounds identified that, when fitted to Eqn. (8), yielded pIC_{50} values more than 4.3, which is equivalent to an IC_{50} value of 50 µM. These data confirm that the small molecules identified from ELT binding studies, to some degree, are inhibitors of GlmU acetyltransferase activity. These data also demonstrate that the

SAR for compounds related to the proposed pharmacophore is restrictive and small changes to the small molecules lead to loss of inhibition.

The most significant GlmU medicinal chemistry effort to identify hits against the *E.coli* orthologue, and subsequently optimise the promising leads, in terms of both potency and physical chemistry properties (optimise the lead compounds to be more drug-like), was carried out by AZ, as discussed in Section 1.4.5. A sub-structure search of the GlmU compound collection identified 59 compounds that were within a Tanimoto similarity of 0.7 of the most potent compound presented (compound 8) in the manuscript by Stokes, et al. in 2012[436]. As with the compounds related to the ELT-identified molecules, potential inhibitors were titrated from 10 mM stocks and incubated with 5 nM GlmU, prior to initiation of acetyl transfer by the addition of 200 µM Ac-CoA and GlcN-1P, followed by detection of CoA-SH using the DTP absorbance assay (Figure 116).

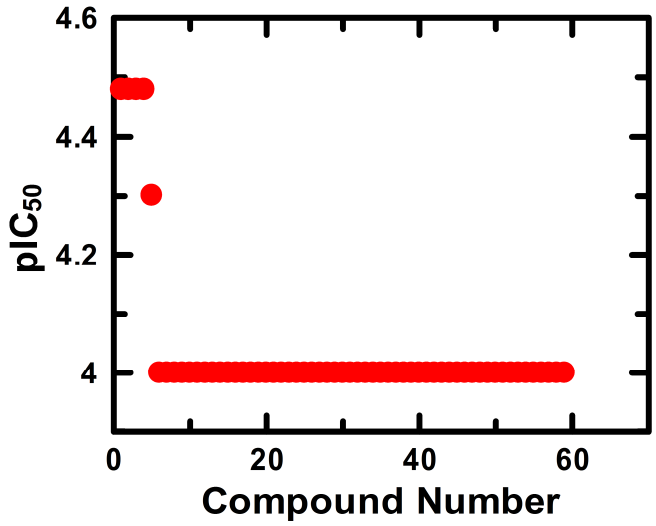


Figure 116. GlmU acetyltransferase screening of compounds related to the most potent AZ *E. coli* GlmU inhibitors.

GlmU acetyltransferase screening of compounds related to the *E. coli* GlmU acetyltransferase inhibitors reported by AZ. The initial rate data were normalised to the microplate control populations using Eqn. (23), before fitting these data to Eqn. (8) to determine the compound pIC_{50} values. The pIC_{50} values for each of the test compounds are represented by (●). The compound number on the x-axis is arbitrary and utilised to demonstrate the spread of pIC_{50} values.

There were 4 compounds identified that, when fitted to Eqn. (8), yielded pIC_{50} values more than 4.3, which is equivalent to an IC_{50} value of 50 µM. These data confirm that the *E. coli* GlmU acetyltransferase inhibitors identified by AZ are poor inhibitors

of *M. Tuberculosis* GlmU acetyltransferase activity. These data demonstrate that the likely mode of inhibition of the AZ inhibitors is specific for *E. coli* and this mechanism does not translate to the *M. tuberculosis* protein. The *E. coli* GlmU acetyltransferase inhibitors identified by AZ were potent against the orthologue from *Haemophilus influenzae* (*H. influenzae*) but 1400-fold selective against the *Streptococcus pneumoniae* (*S. pneumoniae*) enzyme.

Selectivity of GlmU acetyltransferase inhibitors between bacterial orthologues could potentially aid identification of the function of active site residues. To this end, after further searching of the GSK data repositories, a GlmU hit identification project, using the *Staphylococcus aureus* (*S. aureus*) orthologue, was identified. This project was one of many bacterial proteins that were screened against the GSK HTS compound collection and, in the case of both *S. aureus* GlmU enzymatic activities, classified as intractable[348]. This hit identification project, that was terminated in 2000, identified several potential inhibitors of *S. aureus* GlmU acetyltransferase activity, which could provide useful chemical starting points for developing potent inhibitors of *M. tuberculosis* GlmU acetyltransferase. Testing the *S. aureus* GlmU acetyltransferase inhibitors in the *M. tuberculosis* GlmU acetyltransferase assay will allow further understanding of inhibitor selectivity between the acetyltransferase domains of GlmU orthologues. *S. aureus* GlmU acetyltransferase inhibitors were titrated from 10 mM stocks and incubated with 5 nM GlmU, prior to initiation of acetyl transfer by the addition of 200 μM Ac-CoA and GlcN-1P, followed by detection of CoA-SH using the DTP absorbance assay (Figure 117).

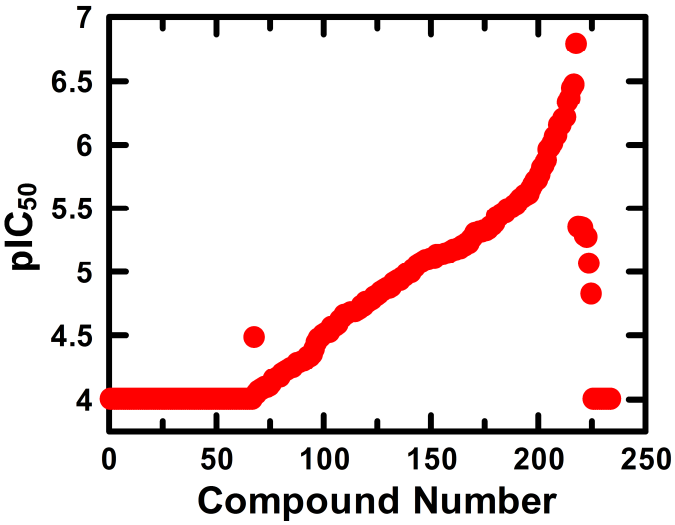


Figure 117. GlmU acetyltransferase screening of *S. aureus* GlmU acetyltransferase inhibitors.

GlmU acetyltransferase screening of GSK compounds identified as inhibitors of the acetyltransferase activity of *S. aureus* GlmU. The initial rate data were normalised to the microplate control populations using Eqn. (23), before fitting these data to Eqn. (8) to determine the compound pIC_{50} values. The pIC_{50} values for each of the test compounds are represented by (●). The compound number on the x-axis is arbitrary and utilised to demonstrate the spread of pIC_{50} values.

There were 234 *S. aureus* GlmU inhibitors tested in the *M. tuberculosis* GlmU acetyltransferase absorbance assay and, after fitting the dose response data to Eqn. (8), 157 compounds had pIC_{50} values greater than 4 (IC_{50} = 100 µM). Several compounds tested had pIC_{50} values greater than 6 and the most potent inhibitor yielded an IC_{50} value of 165 nM. These compounds were previously characterised as pan-GlmU inhibitors (data not shown) and could potentially provide useful tools to better understand the essentiality of the acetyltransferase activity in live *M. tuberculosis* assays.

Finally, a significant number of compounds were screened in dose response experiments from various sources, which primarily included hits and sub-structure searches based on these small molecules, from the TB and robustness single concentration sets. These compounds were titrated from 10 mM stocks and incubated with 5 nM GlmU, prior to initiation of acetyl transfer by the addition of 200 µM Ac-CoA and GlcN-1P, followed by detection of CoA-SH using the DTP absorbance assay (Figure 118).

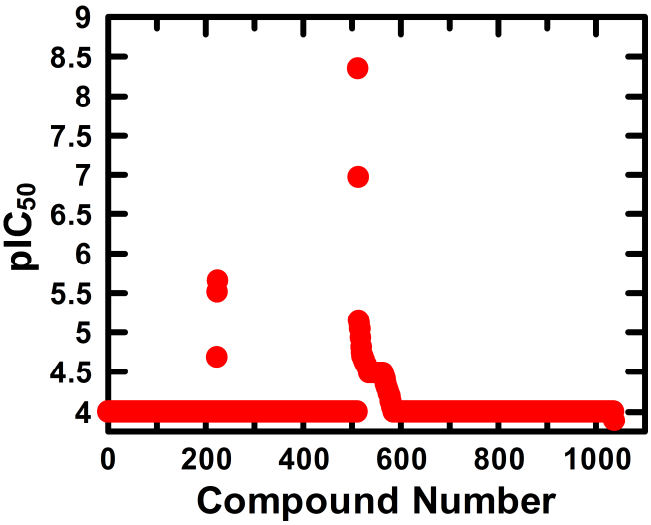


Figure 118. GlmU acetyltransferase dose response screening of hits identified from single concentration screening.

GlmU acetyltransferase dose response screening of hits and compounds related to these hits, identified from single concentration screening and subsequent sub-structure searching. The initial rate data were normalised to the microplate control populations using Eqn. (23), before fitting these data to Eqn. (8) to determine the compound pIC_{50} values. The pIC_{50} values for each of the test compounds are represented by (●). The compound number on the x-axis is arbitrary and utilised to demonstrate the spread of pIC_{50} values.

A total of 1039 compounds were screened in dose response experiments against the GlmU acetyltransferase activity, which identified 25 potential inhibitors with pIC_{50} values greater than 4. The most potent inhibitor identified by these dose response studies had an IC_{50} value of 4.5 nM.

Most of the potential inhibitors identified in these dose response studies need to be further qualified to confirm that the inhibitory responses observed in the acetyltransferase assay are not due to interference mechanisms.

5.6 Knowledge-based inhibitor characterisation

In addition to the inhibitor identification efforts, utilising small molecule screening, an alternative approach was investigated using knowledge-based, rational inhibitor design. To initiate the knowledge-based design effort, GlcN-1P was identified as the starting point for chemistry. The chemistry plan involved modifying the amine group of GlcN-1P, utilising a variety of groups added to block acetylation by GlmU, with the

eventual aim of identifying a non-natural, dead-end inhibitor that would trap the acetyltransferase ternary complex. There were six GlcN-1P analogues synthesised and purified, which are presented in Figure 119.

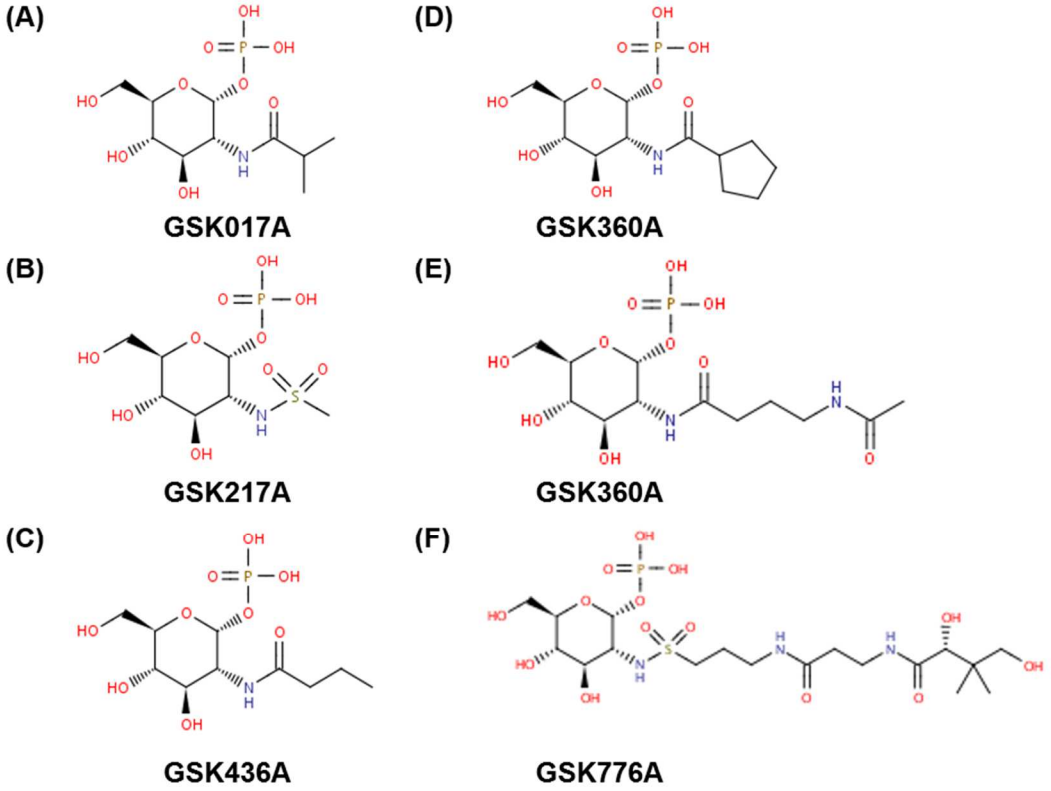


Figure 119. Knowledge-based, compound design of non-natural GlcN-1P analogues.

GlcN-1P analogues designed and synthesised to potentially act as dead-end inhibitors of GlmU acetyltransferase activity. These molecules were prepared by Ben Whitehurst and Rob Young, Novel Chemical Entity Chemistry, GSK.

The first experiment to validate the GlcN-1P analogues was to ascertain that these compounds were not able to replace GlcN-1P as a substrate from GlmU catalysed acetyl transfer. To achieve this, the compounds presented in Figure 119 were titrated in the presence of a fixed concentration of GlmU, prior to initiation of acetyl transfer by the addition of a concentration of Ac-CoA equivalent to 10-fold of K_m (Table 29).

Table 29. GlmU acetyltransferase steady state parameters for non-natural GlcN-1P analogues [a].

GSK compound number	k_{cat} (s^{-1})	K_m (μM)
GlcN-1P	35.3 ± 0.9	199.8 ± 7.2
GSK017A	18.4 ± 2.1	4117.7 ± 508
GSK217A	N.D.	N.D.
GSK436A	1.9 ± 0.3	8824.9 ± 2035.3
GSK930A	N.D.	N.D.
GSK439A	4.2± 0.5	7675.1 ± 1241.5
GSK776A	N.D.	N.D.

[a] Experiments carried out using 10 nM GlmU. Compounds or GlcN-1P, titrated from 6.25 mM and then incubated with GlmU for 15 mins at 25 °C, prior to addition of 1.25 mM Ac-CoA and monitoring CoA-SH formation using 400 μM DTP. Data are a mean of three replicates. Linear initial rates were fitted for each condition and the data were subsequently fitted to Eqn. (1).

The GlcN-1P experiment was carried out as a control concurrently to assessment of the GlcN-1P analogues and the subsequent data generated was as expected. There were some unexpected results from the steady state experiment detailed in Table 29, where when titrating one analogue, GSK017A, a significant increase in k_{cat} (to 18.4 s^{-1}) was observed. Furthermore, modest acetyltransferase activity was monitored when titrating two additional analogues, GSK436A and GSK439A, where the apparent k_{cat} was observed as 1.9 and 4.2 s^{-1}, respectively. The apparent K_m values for the GlcN-1P analogues, were at least 20-fold weaker than the endogenous substrate, which is indicative of these analogues being poor substrates. The structure and purity of all the GlcN-1P analogues synthesised were confirmed by both LC-MS and NMR**,** so given the modifications on the primary amine, GlmU catalysed acetyl transfer was not expected. This was true for three of the GlcN-1P analogues tested; GSK217A, GSK930A and GSK776A, where no acetyl transfer was observed at any concentration of titrant. A possible explanation for the unexpected data observed with GSK017A, GSK436A and GSK439A was the presence of

contaminating concentrations of GlcN-1P, which may have been used as a synthesis intermediate. Subsequently, GSK017A, GSK436A and GSK439A were re-purified to ensure that the chance of any slight GlcN-1P contamination of the analogues was further diminished.

The GlcN-1P analogues were then assessed as inhibitors of GlmU acetyltransferase activity in the presence of the natural substrates. In order to carry out this study, each of the GlcN-1P analogues were titrated from 1 mM and then incubated with GlmU for 15 mins at 25 °C, prior to addition of 200 M Ac-CoA and 100 M GlcN-1P to initiate acetyl transfer and 400 M to monitor CoA-SH formation (Table 30).

Table 30. GlmU acetyltransferase competition studies using non-natural GlcN-1P analogues.

GSK compound number	IC_{50} (µM)	Maximum inhibition (v (s^{-1})) (Concentration of competitor (µM))
GSK017A	30.6 ± 4.2	26.6 ± 1.5 (1000)
GSK217A	40.6 ± 8.6	25.1 ± 2.5 (1000)
GSK436A	71.9 ± 17.1	26.1 ± 3.1 (1000)
GSK930A	91.8 ± 23.5	28.3 ± 3.5 (1000)
GSK439A	60.7 ± 8.9	27.4 ± 1.9 (1000)
GSK776A	N.D.	N.D.

[a] Experiments carried out using 10 nM GlmU. Compounds titrated from 1 mM and then incubated with GlmU for 15 mins at 25 °C, prior to addition of 250 µM Ac-CoA and 125 M GlcN-1P to initiate the reaction and monitoring of CoA-SH formation using 400 µM DTP. Data are a mean of three replicates. Linear initial rates were fitted for each condition and the data were subsequently fitted to Eqn. (8).

The data in Table 30 demonstrate that five of the six GlcN-1P analogues inhibited GlmU acetyltransferase activity with IC_{50} values in the range of 30.6 to 91.8 µM. The compound that did not inhibit acetyltransferase activity was GSK776A, which is the least like the native substrate of the analogues synthesised. To confirm these observations and better define the maximum and minimum asymptotes of the dose response plots, the inhibition studies were repeated. In the repeat study, the GlcN-

1P analogues were titrated from 10 mM and then incubated with GlmU, before initiation of acetyltransferase activity by the addition of GlcN-1P and Ac-CoA (Figure 120).

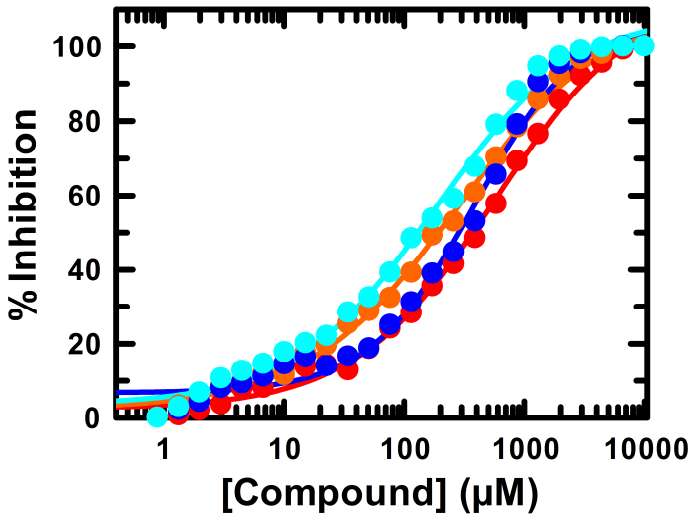


Figure 120. GlmU acetyltransferase, GlcN-1P analogue dose response plots.

Titrations of GlcN-1P analogues from 10 mM, incubated with 2.5 nM GlmU, prior to initiation of acetyl transfer with 250 µM Ac-CoA and 125 µM GlcN-1P. CoA-SH formation was detected by 400 µM DTP. The compound titration data presented are repeats of two of the analogues; GSK439A (●, ●) and GSK930A (●, ●). The data are a mean of three replicates. Linear initial rates were fitted for each condition, normalised to the on-microplate controls using Eqn. (23) and subsequently, the data were fitted to Eqn. (8) to determine IC_{50} values.

The GlcN-1P analogue dose response data presented in Figure 120 typifies the inhibition data generated with these compounds. Notably, after fitting these data to Eqn. (8) the apparent IC_{50} values for GSK439A and GSK930A, were 421.5 ± 48.2 and 279.1 ± 30.7 µM, respectively. The apparent IC_{50} values determined from these studies were 7- and 3-fold, for GSK439A and GSK930A, respectively, weaker than the original inhibition data, presented in Table 30. A similar trend was observed for both GSK436A and GSK217A, with the observed IC_{50} values 10-fold weaker than the original inhibition data. After multiple re-tests, the original inhibition data determined for GSK017A could not be repeated, while no inhibition was observed from GSK776A.

The inhibitor studies utilising the GlcN-1P analogues were variable and the data were hard to consistently interpret. There were several contributing factors to the variability observed with these inhibition studies, primarily, the lack of consistency

between experiments but also, at the time these experiments were carried out, there was limited understanding of the GlmU acetyltransferase kinetic mechanism.
The data from these studies, in addition to potential future approaches and experiments will be further discussed in the Discussion section.

5.7 Inhibition of *Mycobacterium tuberculosis* growth by confirmed GlmU acetyltransferase inhibitors

To further characterise the validated GlmU acetyltransferase inhibitors identified in previous sections in this chapter, a sub-section of these compounds were tested against *M. tuberculosis* to investigate whether they had any effect on the growth of the live bacteria. The compounds selected for screening against *M. tuberculosis* were GSK920684A and the close analogues; GSK024A, GSK025A and GSK026A. The rationale for selecting these small molecules was twofold; GSK920684A was an original GlmU acetyltransferase hit from screening the TB set and by how this set was selected, a minimum inhibitory concentration (MIC) was likely. In addition, the four compounds selected had been confirmed as GlmU acetyltransferase binders or inhibitors utilising several different GlmU assay approaches. The GlmU acetyltransferase inhibitors were titrated from 50 mM before adding *M. tuberculosis* inoculum and incubating the test microplate for six days at 37 °C. Detection solution was added to the test microplate, which was then subsequently for a further two days at 37 °C before measuring the assay signal to determine the MIC (Table 31).

Table 31. Determination of *M. tuberculosis* MIC values for GlmU acetyltransferase inhibitors.

GSK compound number	Extracellular MIC (µM) [a, b]	Intracellular MIC (µM) [a, b]
GSK920684A	80	24.8 ± 2.1
GSK024A	10	1.6 ± 0.2
GSK025A	35	17.2 ± 1.9
GSK026A	40	21.3 ± 0.7

[a] Experiments carried out in duplicate. [b] MIC values were fitted from assay data were as per Section 2.2.30.

The data presented in Table 31 were carried out by Beatrix Rodriguez, TB Discovery Performance Unit, GSK. The MIC data for GSK920684A and the three close analogues in Table 31 demonstrated that all these inhibitors prevent *M. tuberculosis* growth, both directly by monitoring the effect on the bacteria in culture and also when measuring the effect of the test compounds on *M. tuberculosis*-infected THP-1 cells[336]. The MIC results for both the intra- and extracellular values of GSK920684A and analogues vary in terms of absolute apparent measure, but a rank order is maintained; GSK024A is the most potent compound and GSK920684A is the weakest inhibitor.

The MIC's determined for the test compounds in Table 31 are like the GlmU acetyltransferase apparent IC_{50} values. These data confirm that inhibition of *M. tuberculosis* growth by GSK920684A and related compounds could be due to inhibition of GlmU acetyltransferase activity.

Chapter 6. Discussion

6.1 Acetyltransferase kinetic mechanism

Initially, *M. tuberculosis* GlmU was expressed and purified to homogeneity from *E. coli,* which yielded 2.96 mgs of His_6-TEV-GlmU from every 1 gram of cell pellet. The monomer mass of His_6-TEV-GlmU was confirmed by LC-ESI-TOF-MS, which corresponded to the expected molecular weight, minus the *N*-terminal methionine residue. Further analysis of the purified protein, using peptide-mass fingerprinting, confirmed the identity as His_6-TEV-GlmU. Expression and purification of the *C*-terminal TEV-His_6 construct of GlmU identified that there was significant post-translation cleavage. These data suggest that any modification of GlmU *C*-terminal is not well tolerated, leading to proteolysis by endogenous *E. coli* hydrolases. The 30-residue, *C*-terminal extension of *M. tuberculosis* GlmU is not present in *E. coli* and *S. pneumoniae* and is comprised of two α-helices, which are linked by a series of β-turns[215]. The *C*-terminal extension forms part of the acetyltransferase active site, contributing Trp460 and Lys464, which interact with, and stabilise, Ac-CoA. Mutation of Trp460 to Ala has previously been found to significantly compromise acetyltransferase activity. In the same study as the aforementioned W460A SDM, addition of seven residues to the *C*-terminal led to ablation of acetyltransferase activity[245]. Acetyltransferase activity experiments with *E. coli* GlmU identified that constructs with a *C*-terminal hexahistidine tag are 12-fold less active that the native enzyme[437]. These studies demonstrate the importance of the GlmU *C*-terminal extension for acetyltransferase activity and highlight that addition to this region are not likely to be well tolerated.

Further characterisation of His_6-TEV-GlmU, using both CD and CD T_m confirmed that the enzyme had secondary protein structure that could be denatured upon application of a thermal gradient. The data generated using CD to monitor thermal unfolding of GlmU secondary protein structure was confirmed using SYPRO Orange DSF. The T_m values determined using these two biophysical techniques are different, with the DSF values 5 °C higher than the CD determined melting temperature. The main difference between the conditions for these two experiments is that 10 mM $MgCl_2$ was included in the assay buffer for the DSF studies, whereas, due to the known interference of high concentrations of Cl^- ions, it was omitted from

the CD study[438]. Divalent metal ions, such as Mg^{2+} and Co^{2+}, have previously been observed, using X-ray crystallography studies, bound to GlmU in both the *N*- and *C*-terminal domains (Table 7). The absence of divalent metal from the assay buffer for the CD thermal denaturation experiment could lead to a reduction in GlmU T_m, especially if the role of the metal ion is to stabilise the protein structure.

The acetyltransferase activity of GlmU was first evaluated using the thiol detection reagents DTP and DTNB, to kinetically monitor accumulation of CoA-SH by a change in absorbance at either 324 nm or 412 nm, respectively. The two thiol detection reagents were found to have very similar sensitivities to CoA-SH concentration in assay buffer both as a product titration and when detecting product release as a function of acetyltransferase activity. The thiol detection reagent selected to further study GlmU acetyltransferase was DTP, primarily due to the suitability of this reagent for pH-rate studies at pH conditions below 7.3[420]. The use of thiol-reactive detection reagents to kinetically monitor GlmU catalysed acetyl transfer to GlcN-1P, demonstrated that, there are no catalytically important Cys residues. In addition, pre-incubation of GlmU with an excess of thiol modifying reagent, such as NEM, did not have any effect on the acetyltransferase rate. There were some reports that either pre-treatment with thiol-modification agents or mutation of Cys residues to Ala, inactivated *E. coli* GlmU acetyltransferase activity[208,437], however, the data obtained with the *M. tuberculosis* GlmU homologue in this report, does not support these observations and agree with previous findings with the same enzyme[215].

The evaluation of assays based on CoA-SH production identified that, regardless of thiol detection reagent used, initial non-linear acetyltransferase activity (i.e. lag phase) was observed. This non-linear acetyltransferase activity was both enzyme and substrate concentration independent but $MgCl_2$ concentration and incubation-time dependent. The data presented in Figure 29 demonstrate that the acetyltransferase lag phase can be alleviated by incubating GlmU with 10 mM $MgCl_2$ for an extended period at 4 °C. To confirm these observations, the opposite approach was taken, whereby GlmU was incubated with the metal collating agent EDTA for 12 hrs at 4 °C, prior to dialysis to remove EDTA. Subsequently, the acetyltransferase activity of the EDTA-treated preparation of GlmU was assayed in the absence of any $MgCl_2$, which far accentuated the initial non-linear kinetics that had previously been observed (Figure 30). Initial non-linear kinetics for enzymes is not without precedent in the literature and in some cases a requirement for a cofactor,

such as a divalent metal ion has been found to alleviate the observed initial lag phase. In a study with the *M. tuberculosis* enzyme succinic semialdehyde dehydrogenase (GabD1) the researchers identified an initial non-linear phase when monitoring the initial velocity of the $NAD(P)^+$-coupled oxidation of succinic semialdehyde (SSA) to succinate. This non-linear phase was eradicated only when the activity of GabD1 was monitored in the presence of Mg^{2+}, with Co^{2+}, Ni^{2+}, Ca^{2+} and Mn^{2+} demonstrated to have a similar effect. Subsequent studies defined the GabD1 Mg^{2+} $K_{act,}$ value to be 0.3 ± 0.07 mM and revealed a 4-fold decrease in the K_m value for SSA and a 6-fold decrease in the K_m value for $NADP^{+}$[439]. Another example of the activation of a *M. tuberculosis* enzyme was reported for 1-Deoxy-d-xylulose-5-phosphate (DXP) isomeroreductase, which catalyses the isomerisation and reduction, in a nicotinamide adenine dinucleotide phosphate- (NADPH-) dependent manner of DXP to generate 2-C-methylerythritol 4-phosphate (MEP). DXP isomeroreductase was found to be metal ion-activated, displaying specificity for Co^{2+}, in preference to either Mn^{2+} or Mg^{2+}, with the later K_{act} determined to be 1.2 mM, 1000-fold less potent than the preferred cation. However, the authors of the report reason that Mg^{2+} may still be the physiological activating metal ion, as the reported bacterial free intracellular concentration of Mg^{2+} (in *E. coli* of roughly 1−2 mM)[440] agreed with the measured DXP isomeroreductase Mg^{2+} K_{act}[441]. The bacterial intracellular concentrations for Mn^{2+} (0.01μM)[442] and Co^{2+} (0.001μM)[443] are considerably smaller than the corresponding DXP isomeroreductase cation K_{act} values.

These initial observations of the requirement of Mg^{2+} for optimal GlmU acetyltransferase led to a more thorough investigation to fully characterise this divalent metal dependence. This requirement has not been previously characterised in depth with any GlmU homologue, as divalent metals are typically not required cofactors in acetyl transfer reactions. The majority of GlmU homologue studies published to date have used a fixed concentration of Mg^{2+}, in the range of 0.5 to 10 mM[208,209,215,245], presumably based on the early work of Mengin-Lecreulx and van Heijenoort[208]. Experiments utilising EDTA pre-treatments of GlmU not only exacerbates the initial non-linear acetyltransferase phase but, in the absence of divalent metal ion(s) in assay buffer, $k_{cat,\ acetyltransferase}$ was reduced by 5 – 7-fold (Figure 45). The EDTA pre-treatment and subsequent dialysis of GlmU decreased the DSF determined T_m by 2 °C, which indicated that the divalent metal ion(s) had a stabilising effect. The $K_{act,\ Mg2+}$ is 1.5 ± 0.2 mM and the optimal concentration of Mg^{2+}

for maximal GlmU acetyltransferase activity is 10 mM, which is between the 5 and 15 mM Mg^{2+} concentrations determined for the *Y. pestis*[224] and *B. subtilis* homologues[222], respectively. The preparation and storage of GlmU in the continuous presence of 10 mM $MgCl_2$ led to the subsequent eradication of non-linear acetyltransferase activity, which demonstrated the importance of Mg^{2+} for optimal catalysis.

To better understand whether the requirement of GlmU for optimal acetyltransferase activity is specific to Mg^{2+}, several divalent metal ions, including Ca^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+} and Zn^{2+} were investigated. The divalent metal ions could be broadly divided into two groups, primarily based on activation of GlmU acetyltransferase activity and stabilising effect on T_m. The metal ions that activated optimal, linear acetyltransferase activity, in addition to stabilisation of the GlmU T_m, were Mg^{2+}, Ca^{2+} and Mn^{2+}. Indeed, the GlmU acetyltransferase K_{act} for both Ca^{2+} and Mn^{2+} were both lower than the apparent value determined using Mg^{2+}. The effect of these divalent metal ions on the GlmU T_m were contrasting; Mg^{2+} stabilised by 1.2 °C, Mn^{2+} by 6.4 °C, while Ca^{2+} stabilised by a modest 0.3 °C. The GlmU $K_{act,\ acetyltransferase}$ and T_m values demonstrate that Mn^{2+} is the most optimal divalent metal ion for optimal acetyltransferase activity and stabilisation of GlmU, followed by Mg^{2+} and Ca^{2+}. There are a number of reports that GlmU binds Mg^{2+}, Ca^{2+} and Mn^{2+}, primarily, either in the pyrophosphorylase active site[212,246,247], or at the crystallographic axis of symmetry between the three, *C*-terminal L*β*H helices involved in forming the homotrimer[214,220]. Two octahedrally coordinated Mg^{2+} ions were detected on the *H. influenzae* GlmU threefold crystallographic axis, the location of which corresponds to Mg^{2+} and Co^{2+} ions identified in the crystal structure of *E. coli* GlmU in complex with UDP-GlcNAc and Ac-CoA (pdb 2OI7)[220]. The two Mg^{2+} ions in the *H. influenzae* GlmU crystal structure complexes were coordinated by three H_2O molecules and three Asp406-OD atoms, one from each monomer of the homotrimer (pdb 2V0J)[214]. The *M. tuberculosis* GlmU ternary complex crystal structure determined for this project (pdb 6GE9) has one Mg^{2+} bound at the threefold crystallographic axis, coordinated by three Asp417-OD and Ser392-OG atoms[444]. The majority of liganded *M. tuberculosis* GlmU structures in the pdb have two divalent metal ions, either Mg^{2+}, Co^{2+} or a combination of the two, coordinated in the same position by Asp417-OD atoms.

The divalent metal ions that destabilised GlmU, prior to determination of T_m using DSF, were Co^{2+}, Ni^{2+} and Zn^{2+}. These metal ions appear to cause rapid protein aggregation, prior to application of the thermal gradient. This observation is due to the decrease in total maximal fluorescence of GlmU in the presence of either Co^{2+}, Ni^{2+} and Zn^{2+}, when compared to the no metal ion control. The maximal fluorescence in a SYPRO Orange DSF experiment is protein concentration dependent and a decrease in fluorescence is likely due to formation of high order protein aggregates. These protein aggregates occlude the hydrophobic patches of the globular protein core, which are typically exposed and bound by SYPRO Orange upon application of a thermal gradient in a DSF experiment. Aggregation of the protein, in these cases caused by the divalent metal ions, prior to addition of the SYPRO Orange, would bury core hydrophobic patches and prevent a subsequent increase in fluorescence. This divalent metal ion induced protein aggregate formation, would reduce the total concentration of correctly folded GlmU and appear as though less protein had been added used in the DSF assay. DSF thermal stability experiments are very sensitive, so, despite metal ion induced aggregation reducing the total concentration of functional protein, it was still possible to monitor temperature induced unfolding of GlmU. The GlmU T_m was stabilised by increasing concentrations of both Co^{2+} and Ni^{2+} up to maximums of 3 °C and 4 °C, respectively.

Activation by divalent metal ions, specifically Mg^{2+}, Ca^{2+} and Mn^{2+}, has previously been reported for the related *M. tuberculosis* enzyme tetrahydrodipicolinate *N*-succinyltransferase (DapD), which is a trimer in solution and a member of the LβH acyltransferase superfamily[89]. Indeed, the divalent metal binding site is found in a similar position, at the trimer, three-fold axis of symmetry in both DapD and GlmU. The Mg^{2+} ions in the DapD structure are coordinated by the three Asp166-OD atoms and neutralises the negatively charged rings formed by the side chains of Asp164, Asp166 and Glu183 and their respective symmetry mates. DSF studies with *M. tuberculosis* DapD identified that Mg^{2+} stabilised the T_m of the homotrimer[104]. THDP *N*-succinyltransferase activity catalysed by DapD exhibits the same non-linear acyltransferase activity in the absence of divalent metal ions in the assay buffer[104]. The addition of Mg^{2+}, Ca^{2+} and Mn^{2+} resulted in moderate increases in DapD catalysed THDP *N*-succinyltransferase activity, whereas Zn^{2+} or Co^{2+} and inhibited the reaction rate. The inhibition of THDP *N*-succinyltransferase activity by Zn^{2+} or Co^{2+} was also observed in studies with *E. coli* DapD[445]. Additional DapD X-ray

crystallography studies identified that soaking with either excess Zn^{2+} or Co^{2+} led to replacement of the Mg^{2+} at the trimer, three-fold axis of symmetry, coordinated by the same residues. Schuldt, *et al.* propose that the central H_2O filled channel that runs along the homotrimer axis of symmetry allows divalent metal ion access to binding site on DapD[104]. GlmU and DapD are unique amongst the hexapeptide acyltransferases, both enzymes have conformationally mobile *C*-terminal tails that form the active sites upon substrate binding[212]. The structural evidence for a unique substrate-induced conformational change of the *C*-terminal tails of GlmU and DapD, could potentially be stabilised by binding of divalent metal ions at the three-fold trimer axes. Taking these previous observations into account, along with the findings from this project, it seems plausible to conclude that the divalent metal binding to the *C*-terminal of GlmU, plays a structural, stabilising role. Divalent metal ion binding leads to activation of GlmU acetyltransferase activity through a non-catalytic mechanism, potentially by ensuring that the enzyme is in the optimal conformation for catalysis of acetyl transfer. The binding of divalent metal ions to GlmU may provide a novel regulatory mechanism that modulates acetyltransferase activity under specific cellular conditions. Importantly, no evidence for changes in oligomeric state have been observed (data not shown) and GlmU appears to be a trimer under all experimental conditions tested.

The measured steady-state kinetic parameters, presented in Sections 3.6 and 3.7, were of a similar magnitude to those previously determined for *M. tuberculosis* GlmU, as well as other orthologues, specifically $K_{m,\ Ac\text{-}CoA}$ and $K_{m,\ GlcN\text{-}1P}$, which were reported to be in the range of 200 – 400 μM and 60 – 360 μM, respectively (Table 8). Furthermore, the $K_{m,\ UTP}$ and $K_{m,\ GlcNAc\text{-}1P}$ parameters for the uridylyltransferase activity are within a similar magnitude to the reported values, which are 10 – 70 μM and 20 – 110 μM, respectively (Table 8). The reported k_{cat} values for both acetyltransferase and uridylyltransferase activities appear to be more variable, in the ranges of 50 – 1500 s^{-1} and 1 – 350 s^{-1}, respectively (Table 8). This variation in k_{cat} may be caused by the use of discontinuous assays in the majority of L*β*H acyltransferase and pyrophosphorylase kinetic characterisation studies, which inherently lead to greater variabilities and error. Nonetheless, the measured GlmU $k_{cat,\ acetyltransferase}$ value was 28 s^{-1}, which is less than two-fold lower than the reported range. Furthermore, the measured GlmU $k_{cat,\ uridylyltransferase}$ value was110 s^{-1}, which is within the range of the reported values (Table 8). Interestingly, the majority of the

GlmU steady-state kinetic parameter studies to date have reported that the rate-limiting step, and therefore the slower of the two reactions was the uridylyltransferase activity. This is contrary to the findings of this study, which identified that the acetyltransferase activity was the slower of the two enzymatic functions catalysed by GlmU. Early GlmU domain dissection studies on the *E. coli* orthologue suggested that GlcNAc-1P was released from the acetyltransferase domain and allowed to accumulate to K_m concentrations, prior to uridylyltransferase[209]. However, the results from this study, indicate that the opposite is likely, whereby GlcNAc-1P is produced, released and then immediately catalysed by the uridylyltransferase domain to form UDP-GlcNAc. If this is the case and GlmU is the sole GlcNAc-1P consuming activity present in *M. tuberculosis*, a finite concentration will be observable *in vitro*.

The acetyltransferase kinetic mechanism of all GlmU orthologous acetyltransferase activities have been poorly characterised, with no initial velocity, product or dead-end inhibition pattern-type studies carried out. For example, product inhibition of the acetyltransferase activity by CoA-SH, and the lack of inhibition by GlcNAc-1P, have not been described to date (Table 14). Initial velocity patterns obtained by varying GlcN-1P concentrations at fixed concentrations of Ac-CoA were intersecting, which demonstrated that GlmU acetyltransferase activity requires the formation of a ternary complex prior to catalysis. This result rules out a ping-pong mechanism, which would typically lead to the formation of an acetyl-enzyme intermediate. Further studies, using intact mass, protein LC-ESI-TOF-MS were unable to identify an acyl-GlmU intermediate, confirming that a ping-pong mechanism was unlikely (Figure 50 and Figure 51).

The substrate specificity for GlmU acetyltransferase activity was probed with a variety of analogues of Ac-CoA and GlcN-1P. As previously reported for *E. coli* GlmU, there is some flexibility in the length of the acyl moiety that can be transferred from CoA-SH to the amine group of GlcN-1P; with Pro-CoA and Suc-CoA identified as substrates with roughly 10 and 100-fold less efficiency when compared to Ac-CoA (k_{cat}/K_m, Table 13 and Table 14)[209]. The substrate preference of the acetyltransferase active site appears to be limited to acyl-chain lengths of 3-carbons, as when the 4-carbon acyl-CoA analogue (Bu-CoA) was investigated no CoA-SH formation could be detected. However, even though Bu-CoA was no longer a substrate it was still able to bind GlmU and was shown to inhibit GlmU acetyltransferase activity. The acyl-CoAs, AAc-CoA and Suc-CoA were both

identified as substrates (significantly lower k_{cat}/K_m, Table 13) but are of similar thioester chain length to Bu-CoA. The identification of these acyl-CoAs as GlmU substrates could be explained by acyl-thioester hydrolysis, which, due to the thiol detection methodology used to monitor the acetyltransferase reaction, would still appear as product formation.

The most plausible alternative physiological acyl substrate identified for GlmU acyltransferase activity was Pro-CoA, primarily due to the low K_m, which was 4-fold lower than $K_{m,\ Ac\text{-}CoA}$, and the subsequent calculated catalytic efficiency. Although the $K_{m,\ Pro\text{-}CoA}$ was low, the GlmU acyltransferase catalytic efficiency, utilising Pro-CoA as a substrate, was determined as 100-fold worse than the acetyltransferase k_{cat}/K_m. This is likely to be due to the increased length of the acyl-chain of Pro-CoA, which has the direct effect of increasing the acyltransferase $K_{m,\ GlcN\text{-}1P}$ by 3.4-fold. The increased acyl-chain of Pro-CoA also reduces the acyltransferase k_{cat} by 56-fold, when compared to acetyltransferase k_{cat}, which is likely primarily due to the disruption of GlcN-1P binding. However, it was possible to carry out acyltransferase initial velocity pattern studies by varying GlcN-1P concentrations at fixed concentrations of Pro-CoA, which were intersecting, confirming that propionyl transfer activity, like acetyl transfer, requires the formation of a ternary complex prior to catalysis.

The GlmU acetyltransferase substrate specificity for amino sugar phosphate is much more restricted when compared to the previously discussed Ac-CoA specificity. The only other amino sugar phosphate identified as an alternative acetyltransferase substrate was UDP-GlcN, which had previously identified in studies with the *E. coli* orthologue[209,210]. The identification of UDP-GlcN as an alternative GlmU acetyltransferase substrate, albeit a poor one ($k_{cat}/K_{m,\ UDP\text{-}GlcN} = 2.2 \times 10^2$ in comparison to $k_{cat}/K_{m,\ GlcN\text{-}1P} = 9.7 \pm \times 10^4$), raises the possibility of an alternative order of GlmU catalysed reactions (Scheme 11)[209]. However, it was not possible to confirm that GlcN-1P could be utilised as a substrate for uridylyl-transfer by the *M. tuberculosis* GlmU, in contrast to the *E. coli* orthologue studies (that conflicted with previous studies by Gehring, *et al.*)[210]. These data suggest that GlmU may be able to catalyse formation of UDP-GlcNAc from UDP-GlcN *in vivo* but this is unlikely to be the preferred route and probably only occurs in response to specific physiological conditions. These results confirm that the canonical order of GlmU reactions is

preferred route for formation of UDP-GlcNAc for the *M. tuberculosis* orthologue (Scheme 11).

The other GlcN-1P analogues investigated demonstrated the requirements required for acetyltransferase active site binding. Glucosamine and GlcN-6P were not utilised as substrates, which highlights the importance of a phosphate group at the C-1 position of the amino sugar. In contrast to findings in a previous study characterising the substrate specificity of the *E. coli* orthologue, galactose-derived amino sugar analogues (C-4 epimers of glucose) are not utilised as substrates for the acetyltransferase activity of *M. tuberculosis* GlmU. These results highlight that *M. tuberculosis* GlmU has a stronger preference for the glucose epimer when compared to the *E. coli* orthologue.

GlmU acetyltransferase substrate analogues that were not substrates were further investigated as potential inhibitors. The inhibition of acetyltransferase activity by GlcN-1P analogues was observed only with Glc-1P (less than 10 mM IC_{50}), which confirms the binding specificity for the glucose C-4 epimer. Inhibition was not observed with either GlcNAc-1P or GlcNAc, which demonstrates that acetylated products have a low affinity for the GlcN-1P binding site. To further understand inhibition by Glc-1P, X-ray crystallography studies, in the presence of Ac-CoA were carried out, which identified that it was possible to overlay the position and orientation of Glc-1P with GlcN-1P (using a previously solved alternate ternary complex, pdb 3ST8, Figure 67). Importantly, the GlcN-1P amine and the Glc-1P hydroxyl both occupy the same position in the ternary complex structure, as do the rest of the molecules. The hydroxyl in position C-2 of Glc-1P is within hydrogen bond distance to His374, which suggest that this residue could serve as the general base that deprotonates the GlcN-1P amine prior to nucleophilic attack on Ac-CoA. A similar approach to study the ternary Michaelis complex of *Aspergillus fumigatus* GlcN-6P *N*-acetyltransferase was carried out using Glc-6P as a dead-end substrate in place of GlcN-6P, which provided evidence to aid inference of the likely catalytic mechanism[446].

Further *M. tuberculosis* GlmU acetyltransferase dead-end inhibition studies were carried out using the Ac-CoA analogues not initially identified as substrates. These studies identified dead-end inhibition with the following Ac-CoA analogues; Bu-CoA, IsoBu-CoA, Mlo-CoA and Cro-CoA. The acetyltransferase inhibition observed with these Ac-CoA analogues is likely due to the increase in acyl-thioester chain length,

which will extend from the Ac-CoA binding site, protruding into the GlcN-1P site and subsequently precluding GlcN-1P binding to GlmU. Product inhibition by CoA-SH was relatively potent (when compared to the apparent $K_{m,\ Ac\text{-}CoA}$), whereas no inhibition was observed with GlcNAc-1P, even at concentrations 180-fold higher than $K_{m,\ GlcN\text{-}1P}$. No product-analogue inhibition was observed with dethio-CoA, suggesting that additional interactions with the sulphur of the β-mercaptoethylamine portion of CoA-SH, reported by Jagtap, *et al.*, are required for binding[245].

DSF binding studies facilitated determination of equilibrium constants for the dissociation of Ac-CoA, Bu-CoA, CoA-SH, GlcN-1P and Glc-1P from GlmU. Titrations of Ac-CoA, Bu-CoA and CoA-SH were able to stabilise GlmU, increasing the observed T_m by up to 4 °C and subsequent use of the change in T_m (ΔT_m), as a function of substrate ligand concentration, allowed dissociation constants to be determined. These K_d values were broadly in agreement with the K_m and IC_{50} values that were previously determined for these substrates, products and dead-inhibitors. Dissociation constants for GlcN-1P and Glc-1P from GlmU could not be determined without Ac-CoA or a close analogue present, however, in the presence of high concentrations of a known interacting acyl-CoA, K_d for the complex could be determined. These findings support an ordered bi-bi sequential kinetic mechanism, whereby Ac-CoA or a short acyl-CoA analogue binds to GlmU first, followed by completion of the ternary complex with GlcN-1P association.

To further characterise whether the likely bi-bi sequential kinetic mechanism is either random or ordered, both product and dead-inhibition studies were carried out. The inhibitor patterns obtained using the product CoA-SH are most consistent with a sequential ordered bi-bi kinetic mechanism (Scheme 12). CoA-SH was determined to be a competitive inhibitor when titrated against Ac-CoA and a non-competitive inhibitor when varied against fixed concentrations of GlcN-1P. Crucially, the formation of a dead-end ternary complex using Glc-1P as an inhibitor, which was determined to be uncompetitive when varied against fixed concentrations of Ac-CoA. The uncompetitive inhibition pattern observed with varying concentrations of Glc-1P in the presence of fixed concentrations of Ac-CoA does not change regardless of the concentration of GlcN-1P utilised (Figure 61).

Further dead-end inhibition studies identified two competitive inhibition patterns (Table 15), which were observed with titrations of Glc-1P versus GlcN-1P and Bu-CoA versus Ac-CoA. These competitive inhibitor patterns are characteristic of an

ordered mechanism with the formation of non-productive GlmU:Ac-CoA:Glc-1P and GlmU:Bu-CoA complexes. Bu-CoA likely extends part of its acyl-chain into the GlcN-1P active site, which blocks GlcN-1P from binding therefore appearing as a competitive inhibition. Indeed, this observation is confirmed by mutual exclusivity studies carried out using Bu-CoA and Glc-1P. Analysis of this experiment using the Yonetani-Theorell method confirms that these inhibitors are partially mutually exclusive. An α value of 3 was obtained from this experiment, suggesting some degree of antagonistic binding, which further confirms competition likely due to the extended acyl-thioester chain length of Bu-CoA (Figure 60).

Importantly, no inhibition was observed with GlcNAc-1P against either Ac-CoA (in the presence of varying fixed concentrations of GlcN-1P) or GlcN-1P. This lack of inhibition is diagnostic of an equilibrium ordered bi-bi kinetic mechanism. However, contrary to this observation, CoA-SH inhibition versus GlcN-1P is non-competitive, instead of competitive, as expected for and equilibrium ordered mechanism. A protein conformational change of GlmU induced by Ac-CoA binding but prior to GlcN-1P binding could potentially explain both the initial velocity and inhibitor patterns obtained and demonstrates that the most likely kinetic mechanism is steady-state ordered.

The structure of GlmU with Ac-CoA and Glc-1P bound, provides useful insights into both formation of the catalytic and dead-end inhibited ternary complexes. Inhibition by CoA-SH and Bu-CoA versus Ac-CoA, which in both cases is competitive, can be explained by the conserved recognition and binding of both the adenine ring and pantetheinyl arm of CoA, in combination with the limited interactions with the thioester bound to the β-mercaptoethylamine. It was apparent from the GlmU:Ac-CoA:Glc-1P structure that the binding sites are non-overlapping, hence, for any competitive inhibition to occur between dead-end analogues of the two substrates, there must be an increase in size of one or both molecules to cause a direct steric clash. Superposition of the GlmU:Ac-CoA:Glc-1P structure with previously determined *M. tuberculosis* GlmU:Ac-CoA and apo-structures, clearly highlights that a disordered loop becomes ordered upon binding of either CoA-SH or CoA thioesters. This conformation change is likely to be facilitated through the interaction of the backbone carbonyl of Tyr398, which is located on the mobile loop, and the nitrogen N4P of Ac-CoA (Figure 68). The conformational change and subsequent stabilisation of the mobile loop, succeeding initial Ac-CoA binding, positions Asn397

and Lys403, in addition to Tyr398, in favourable orientations to interact with either Glc-1P or GlcN-1P. Additional stabilisation in the acetyltransferase active site is contributed by carbohydrate-aromatic interactions between the phenolic side chain of Tyr398 and the hexose ring of GlcN-1P[447], which have previously been identified in the *E. coli* orthologue[220]. The specificity of GlmU acetyltransferase active site for binding hexose sugars with a phosphate group at the C-1, is demonstrated by the GlmU:Ac-CoA:Glc-1P structure. The C-1 phosphate interacts with three cationic side chains (Lys403, Arg344 and Lys362) as well as the side chains of Asn397 and Tyr377. These active site groups stabilise binding of GlcN-1P and ensure the correct orientation for efficient catalysis. Mutation of any of these GlmU residues were subsequently shown to decrease the $K_{m,\ GlcN\text{-}1P}$ by at least 4-fold, as well as reduce the acetyltransferase k_{cat} (see data in Section 4.6 and further discussion in Section 6.2).

Further observations from the GlmU:Ac-CoA:Glc-1P structure identified an interaction between the side chain of His374 and Glu370, in addition to stacking of the His imidazole against the phenolic side chain of Tyr377. These interactions promote the most advantageous tautomeric form of the His374 imidazole, primarily ensuring the proton lacking Nε2 is directed towards the C-2 of Glc-1P (the amine of GlcN-1P would occupy this position) in addition to increasing the basicity of the side chain. The analogous interactions have previously been observed in the GlmU:CoA-SH:GlcN-1P:UDP-GlcNAc structure of the *E. coli* GlmU orthologue[220], in addition equivalent enzyme residues in a similar active site configuration have been described for the non-L*β*H acetyltransferase CAT[132,147].

The identity of the active site residues, forming an oxyanion hole, involved in the stabilisation of the transition state remains elusive. The residues identified to be involved in stabilisation of the pantetheinyl arm of the substrate in the GlmU:Ac-CoA:Glc-1P structure were Ser416 and Ala434, which are equivalent to the residues involved in stabilising CoA-SH in the *E. coli* orthologue[220]. However, mutation studies with *M. tuberculosis* enzyme did not confirm that Ser416 was essential for catalysis[245], which highlights that this is an area for further study.

These structural observations in combination with kinetic and binding data are consistent with a steady-state ordered mechanism, depicted in Scheme 12.

6.2 Acetyltransferase chemical mechanism

The ionisation of the groups that may be responsible for substrate binding and catalysis in the *M. tuberculosis* GlmU-catalysed acetyl transfer reaction were investigated by carrying out pH-rate profiles, examining k_{cat}, $k_{cat}/K_{m,\ Ac\text{-}CoA}$ and $k_{cat}/K_{m,\ GlcN\text{-}1P}$ (Section 4.1).

Prior to embarking on the extensive pH-rate profile study, it was necessary to carry out several controls to ensure that GlmU is stable and active across the planned pH range. To characterise stability, DSF studies were carried out to monitor GlmU T_m at pH conditions spanning the planned range. The GlmU T_m varies by 3.4 °C between pH 6 and 10, however, there are typical thermal profiles at each pH condition tested, suggesting that the protein remains folded (Figure 72 and Table 18).

The acetyltransferase activity of GlmU was assessed by monitoring CoA-SH release using the DTP absorbance assay at pH conditions from 6 to 10. This control experiment is complex, as due to the nature of the pH-rate study, the steady-state kinetic parameters are expected to be dependent on the pH of the assay conditions. To partially address this, the activity experiments were carried using either GlmU that had been incubated for a minimal period in each of the pH conditions or exposed to the pH conditions for a two-hour period, prior to initialisation of acetyltransferase activity. It was possible to kinetically monitor the linear formation of CoA-SH using the change in DTP absorbance at each of the pH conditions. Additionally, for each pH condition tested, there was a similar rate of CoA-SH formation, regardless of the length of GlmU incubation prior to initiation of acetyltransferase activity (Figure 73). These control experiments demonstrated that GlmU is both stable and able to catalyse acetyl-transfer at pH conditions ranging from 6 to 10.

The interpretation of pH-rate profiles is complex but can report on several groups involved in binding, catalysis or both. Values for k_{cat}/K_m and k_{cat} are obtained by titrating one substrate at saturating concentrations the co-substrate. The pH dependence of k_{cat}/K_m reflects the protonation state of ionisable group(s) on either the free enzyme, substrate, or both, important for all steps from binding of the varied substrate until the first irreversible step. The pH dependence of k_{cat} reflects the groups on the enzyme required for acetyltransferase activity, in this case, from the ternary complex (GlmU:Ac-CoA:GlcN-1P) until release of the last product.

The $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile was described by a bell-shaped curve that decreases at both low and high pH, which was best fit to an equation for two nonresolvable acidic groups and one basic group. The ionisation constants determined for the $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile were pK_a = 6.6 and pK_a = 8.1, which both appear to be important for binding of Ac-CoA and catalysis. The crystal structure of GlmU with Ac-CoA bound in the acetyltransferase active site reveals that the cofactor is stabilised in an extended conformation that directs the Ac-CoA sulphur atom toward the His374. The acidic ionisable group identified in the $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile can be tentatively assigned to His374, which is proposed to be the general base required for deprotonation of the $-NH_3^+$ group from GlcN-1P. The basic ionisable group (pK_a = 8.1) could correspond to Tyr398 which has been shown to interact through its backbone carbonyl with the oxygen atom of the acetyl group of Ac-CoA. A similar pK_a value has been observed previously in studies on the enzyme serotonin *N*-acetyltranferase[448], which led to assignment of this ionisable group to an active site Tyr residue. The free and liganded structures of GlmU (Figure 68) identified a mobile loop that is only visible in X-ray structures when acyl-CoA or CoA-SH are present. The conformational change of this loop, induced by Ac-CoA binding, brings Tyr398 into hydrogen bonding distance with Ac-CoA, which could lead to further stabilisation, prior to catalysis. Further studies, possibly SDM of Tyr398, are required to accurately assign the basic ionisable group, responsible for the pK_a value of 8.1, involved in Ac-CoA binding and catalysis.

The $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profile was described by bell-shaped curve that decreases at both low and high pH, which was best fit to an equation for two nonresolvable acidic and basic groups. The dissociation constants determined from the $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profile were pK_a = 7.5 and pK_a = 9.1, which are both important for binding of GlcN-1P and catalysis. One of the acidic ionisable groups determined to have pK_a = 7.5 can potentially be attributed to the His374 active site residue, which is in the range of 90% of the His residues identified in pH-rate studies (the vast majority of these measurements fall within the range of pH 5 – 8[449]). The basic ionisable groups in the $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profile were determined to have a pK_a = 9.1 are difficult to attribute to any of the ionisable groups in the acetyltransferase active site. Further studies were carried out to assign the basic groups, which will be discussed in depth later in this section.

The k_{cat} pH-rate profile revealed the presence of four nonresolvable ionisable groups, two with pK_a values of 6.7 and two with pKa values of 9.0. The ionisable groups in the k_{cat} pH-rate profile have previously been identified in the k_{cat}/K_m profiles. The pK_a values of 6.7 were identified in both the k_{cat} and $k_{cat}/K_{m,\ Ac\text{-}CoA}$ profiles and has tentatively been assigned to at least one enzyme residue, which is His374. This assignment was further investigated by SDM of His374 to Ala, which is discussed in more depth later in this section. The pK_a values of 9.0 - 9.1 were identified in both the k_{cat} and $k_{cat}/K_{m,\ GlcN\text{-}1P}$ profiles, which have previously been difficult to assign to any ionisable residue in the acetyltransferase active site.

To further assign pK_a values identified in the pH-rate profile study, ^{1}H NMR spectroscopy studies were carried out to ascertain the pK_a values of two GlcN-1P groups, CH(2)-NH_3^+ and CH(1)-PO_4H_2. The pK_a values determined for CH(2)-NH_3^+ and CH(1)-PO_4H_2 were 8.43 and 8.44, respectively. These values do not directly correspond to the pK_a values determined in the k_{cat} or $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profiles, which could potentially indicate that the binding of GlcN-1P and its deprotonation by His374 are potentially not rate-limiting for acetyl transfer. The inability to assign the ^{1}H NMR spectroscopy determined pK_a values to any of the pH-rate acid dissociation constants led to further SDM studies to further characterise these groups.

Initially, His374 was mutated to Ala to ascertain what effect would be caused by removing the imidazole sidechain, initially, on the stability of GlmU and acetyltransferase activity and subsequently on the pH dependence of the steady-state kinetic parameters for both Ac-CoA and GlcN-1P. The stability of GlmU H374A mutant was assessed by DSF, which identified that the T_m remained unchanged, when compared to the wild type enzyme. Additionally, binding of 1000 µM CoA-SH stabilised GlmU H374A T_m by 1.7 °C, which is the same extent as that for WT enzyme. The acetyltransferase activity of GlmU H374A diverged significantly from the wild type enzyme in assay buffer adjusted to pH 7.5, with CoA-SH formation completely ablated at most concentration tested (Figure 79).

Subsequently, to ascertain the effect of mutating His374 to Ala on acetyltransferase steady-state parameters, pH-rate studies were carried out using GlmU H374A. The $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profile was best fit to an equation for a single ionisable group, which revealed a pK_a = 9.7. This is likely an enzymic group involved in the binding of the phosphate of GlcN-1P, possibly either Tyr377 or Arg344. The $k_{cat,\ GlcN\text{-}1P}$ pH-rate profile revealed that as the pH conditions became more basic the $k_{cat,\ GlcN\text{-}1P}$

increased with a slope of one until a plateau at pH 9. The single ionisable group that needs to be deprotonated under k_{cat} conditions was observed, which was determined to have a pK_a value of 8.2. The $k_{cat,\ GlcN\text{-}1P}$ pK_a value is in close agreement with the pK_a value of 8.4 for the amino group of GlcN-1P, which was previously determined using ^{1}H NMR spectroscopy. This result indicates that deletion of the imidazolium group of His374 leads to a significantly decreased catalytic rate (1000-fold) and an alteration in the rate-limiting step. Catalysis of acetyl transfer by GlmU H374A appears to be dependent on deprotonation of GlcN-1P prior to enzyme binding, which highlights that this chemistry is rate-limiting, as all other wild type identified ionisable groups are no longer observed.

The GlmU H374A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile was a bell-shape that decreases at both low and high pH conditions, which was best fit to an equation for two nonresolvable acidic groups and one basic group. The ionisation constants determined for the GlmU H374A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile were pK_a = 7.7 and pK_a = 9.1, which both seem to be important for binding of Ac-CoA and catalysis. The GlmU H374A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pK_a values are divergent from the wild type $k_{cat}/K_{m,\ Ac\text{-}CoA}$ values, which demonstrates that the mutation of His374 to Ala has altered these constants. The GlmU H374A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pK_a values are both increased by one log unit to make both more basic and, interestingly, the same as the acid dissociation constants determined for the GlmU wild type $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profile. These observations demonstrate that the GlmU H374A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pK_a values are difficult to specifically assign but the groups could potentially be attributed to enzymic groups that are involved in GlcN-1P binding.

The GlmU H374A $k_{cat,\ Ac\text{-}CoA}$ pH-rate profile is in close agreement to the $k_{cat,\ GlcN\text{-}1P}$ profile determined with the mutant enzyme. The acid dissociation constant determined from the GlmU H374A $k_{cat,\ Ac\text{-}CoA}$ pH-rate profile was pK_a = 8.2, which has previously been assigned to the amino group of GlcN-1P.

Subsequently, further SDM studies were carried out to try and assign the pK_a values identified for catalysis of acetyl transfer of wild type GlmU. These studies involved mutation of GlmU acetyltransferase active site residues Arg344, Lys362 and Tyr377 to Ala. The same protein stability and enzyme activity control studies were carried out for the three new SDM proteins, as were implemented for the H374A mutant. These studies identified that none of the mutations shifted the GlmU T_m more than 1.5 °C, while the acetyltransferase activity was reduced for R344A and K362A and

ablated for the Y377A mutant (Table 20, Figure 83). The $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profiles for R344A, K362A and Y377A were all best fit to an equation for a single ionisable group, which revealed pK_a values of 7.4, 7.3 and 7.6, respectively. These pK_a values are like dissociation values that have previously been attributed to the His374 active site residue, which has been tentatively assigned as the acetyltransferase general base. The $k_{cat}/K_{m,\ GlcN\text{-}1P}$ pH-rate profiles for R344A, K362A and Y377A were all missing additional pK_a values in the basic range, which could indicate that one or more of these residues can be potentially assigned to these dissociation constants.

The $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profiles for R344A, K362A and Y377A were all divergent from the wild type enzyme profile. There were similar acid dissociation constants determined from fitting these pH-rate profiles, primarily, two non-resolvable pK_a values of 9.1 or 9.2 were identified from the R344A and K362A mutants. These two non-resolvable basic pK_a values were also identified from the wild type GlmU $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile. Importantly these two non-resolvable basic groups are missing from the GlmU Y377A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile, which indicates that one of the pK_a values of 9.1 can be assigned to Tyr377. The acidic groups identified in the $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profiles diverge from the wild type GlmU (pK_a = 6.4) from 0.5 log, for GlmU K362A, to either pK_a = 7.7 or pK_a = 7.8 for the Y377A or R344A mutants. Furthermore, the equations used to fit these data differed from the wild type enzyme, which was fit to an equation describing two nonresolvable acidic and basic groups, as was the GlmU K362A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile. An equation describing one acidic and two nonresolvable basic groups was the best fit for the GlmU R344A $k_{cat}/K_{m,\ Ac\text{-}CoA}$ pH-rate profile, while an equation describing one acidic group fit the Y377A data optimally. These differences are difficult to reconcile, as the mutations do not remove groups that can be directly attributed to the acid dissociation constants identified. However, the mutations may affect local acetyltransferase active site conditions and modulate the determined pK_a values, all of which are within the reported range of pK_a values for His residues[449].

The GlmU $k_{cat,\ Ac\text{-}CoA}$ and $k_{cat,\ GlcN\text{-}1P}$ pH-rate profiles for R344A, K362A and Y377A are divergent from the wild type k_{cat} profiles, as a decrease in k_{cat} under basic conditions is not observed. These data suggest that either one or a combination of these residues are responsible for the basic groups, which are determined to have pK_a values in the range of 9.2 to 9.6, observed in the wild type GlmU k_{cat} pH-rate

profiles. The dissociation constants for the acidic groups observed in the $k_{cat,\ Ac\text{-}CoA}$ and $k_{cat,\ GlcN\text{-}1P}$ pH-rate profiles for R344A and K362A differ by between 1.3 to 1.1 pK_a log units, respectively. The pK_a values determined for GlmU R344A and K362A $k_{cat,\ GlcN\text{-}1P}$ (pK_a = 8.4 and pK_a = 8.2, respectively) are comparable to the pK_a values determined for Y377A from both k_{cat} pH-rate profiles. These pK_a values are comparable to the pK_a values previously determined for GlcN-1P using ^{1}H NMR spectroscopy.

Notably, the magnitude in k_{cat} differs for the pH-rate profiles for each SDM mutant from wild type GlmU, with a 10-fold reduction observed for GlmU Y377A. There is a difference between the k_{cat} of GlmU R344A and K362A from the two pH-rate profiles, which is indicative of the inability to fully saturate with GlcN-1P when titrating Ac-CoA. The $k_{cat,\ GlcN\text{-}1P}$ pH-rate profiles demonstrate that the k_{cat} values for both GlmU R344A and K362A are likely to be the same as the wild type enzyme. These k_{cat} variations demonstrate that the mutation of Tyr377 to Ala has a deleterious effect on the rate of GlmU acetyl transfer by 10-fold, while the mutations of R344A and K362A do not alter the magnitude of catalysis.

Further SDM studies could be carried out to further understand the ionisable groups identified in the pH-rate profiles, in addition to exploring the amino acids identified from the structural biology studies. The Tyr398 residue, which is found on the mobile loop that changes conformation upon Ac-CoA binding would be a candidate for further study, as would Glu370, the residue that has been shown to interact with the imidazole side chain of His374. The two residues that have been speculated to form the active site oxyanion hole, Ser416 and Ala434, would be additional candidates for SDM studies, as would the two-substrate interacting residues, Asn388 and Asn397. In addition to further SDM studies to investigate more GlmU acetyltransferase active site residues, an alternative mutagenesis strategy could be taken to the one employed for this project. The strategy taken here was to mutate single acetyltransferase active site residues to Ala, which in some cases, for instance with His374, is not a conservative change. A conservative mutation is typically an amino acid replacement in a protein that changes a given amino acid to a different amino acid with similar biochemical properties. An alternative SDM strategy that could be explored is to make more conservative mutations, for instance, mutating Tyr377 or Tyr398 to Phe. In addition, the strategy taken for this project was to mutate one residue at a time to attribute specific functions to the amino acid of interest, such as

involvement in catalysis. The single residue SDM approach could be expanded to change two amino acids at a time, which would provide further information about the interplay between groups in the acetyltransferase active site. A combination of a conservative SDM approach, along with changing more than one residue at a time, may help to further elucidate the residues in the acetyltransferase active site involved in substrate recognition, binding and catalysis.

To assess the contribution of solvent derived protons to the rate-limiting steps of the acetyl transfer catalysed by GlmU solvent kinetic isotope effect studies (SKIEs) were utilised. SKIE studies involve the titration of one substrate in the presence of saturating concentrations of all others, in either 100% H_2O or D_2O, which are subsequently used as comparators to determine the effect of replacing the H atom with a D atom[450]. Prior to proceeding with SKIE experiments, it is important to ensure that the effects observed are not due to the higher solvent viscosity associated with the use of D_2O, when compared to the viscosity of H_2O. In the case of GlmU acetyltransferase activity, the addition of 9% glycerol to the H_2O conditions, which acts as a viscogen to mimic the viscosity of 100% D_2O, had only small effects on $V/K_{GlcN\text{-}1P}$[451].

SKIEs effects on the V profiles encompass all steps from GlmU:Ac-CoA:GlcN-1P and includes all steps up to and including the release of GlcNAc-1P and CoA-SH and the subsequent regeneration of the free enzyme. SKIEs effect on the V/K profiles report on all steps from free GlmU until the first irreversible step (usually assumed to be the chemical step).

These SKIE studies identified highly reproducible SKIEs of 1.6 on both $V_{GlcN\text{-}1P}$ and $V_{Ac\text{-}CoA}$, while a small SKIE of 1.34 on $V/K_{Ac\text{-}CoA}$, was determined. Additional studies were carried out to ascertain the number of protons involved in the chemical step evident in the solvent isotope effects, which are known as proton inventory studies[452]. The data from these experiments were linear, which is indicative of a single proton in flight in the rate-determining step of the reaction, which is likely the origin of the observable SKIE on V.

The chemical step that is partially rate-limiting for acetyl transfer and that is observed during these SKIE studies is likely a solvent-sensitive step, such as the deprotonation of GlcN-1P. The similar magnitude of the SKIEs on V obtained by varying either substrate, at saturating concentration of the co-substrate indicates that deprotonation of GlcN-1P by a general base, likely His374, is partially rate-limiting.

This is likely not a simple single isotope effect, as deprotonation of GlcN-1P the deuterated amine will be a combination of the primary isotope effect of the proton being attracted by the His374 imidazole, in addition to the contribution of the other two alpha deuterons, that remain in the amine. Proton inventory experiments, on ^{D2O}V indicates that a single proton is in flight during the transition state. In contrast, this isotope effect could be originated in subsequent steps leading to formation of free enzyme.

The small SKIE observed on $V/K_{Ac\text{-}CoA}$ and the lack of an effect on $V/K_{GlcN\text{-}1P}$ indicate that D_2O has little effect on the formation of the ternary complex and on chemistry. Together, these small SKIEs indicate that steps other than chemistry are likely rate-limiting for GlmU-catalysed acetyl transfer, particularly under *V/K* conditions.

6.3 Acetyltransferase inhibitor identification

Novel inhibitor identification of the GlmU acetyltransferase activity was carried out to provide chemical tools to chemically validate the essentiality of the enzyme function *in vivo* in addition to potentially providing unique starting points for future drug development projects. Furthermore, these inhibitor identification studies will enable a thorough, independent assessment of the chemical tractability of the GlmU acetyltransferase active site, which will provide confidence for future drug discovery efforts. Finally, these studies will provide a comparative dataset to the published GlmU acetyltransferase activity inhibitor identification studies, for both *M. tuberculosis* and orthologues (see Section 1.4.5).

Prior to commencing small molecule screening of the GlmU acetyltransferase activity, it was necessary to configure several assays to be used to identify and confirm novel inhibitors. The primary assay used for initial screening of all small molecule was the DTP thiol detection absorbance assay, however, further optimisation experiments were required to validate its use. These experiments involved ensuring that the appropriate concentrations of both substrates were utilised, optimisation of the DTP concentration and determination of the linearity of GlmU catalysed acetyl transfer activity. The combination of the concentration of reagents used in the acetyltransferase activity DTP absorbance assay, were critical. It was necessary to achieve a signal window of sufficient magnitude to enable determination of a Z' factor value that was in excess of 0.4[453], while still maintaining an assay sensitive to

inhibition and ensuring that the turnover of Ac-CoA does not exceed 50%. The validation experiments confirmed the optimal conditions for the DTP thiol detection absorbance assay, which were the use of 200 µM Ac-CoA and GlcN-1P, 400 µM DTP. The concentration of GlmU selected for use during small molecule screening studies, was 5 nM, which gave a linear response for a sufficient time to achieve good assay signal window, prior to turnover of 50% Ac-CoA.

The GlmU acetyltransferase activity DTP absorbance assay was evaluated for single concentration small molecule hit identification by screening two GSK compound training collections, known as the Robustness and Nuisance sets. These sets are typically used to assess how prone the test assay is to random and known false positive compounds, respectively. The Robustness set is comprised of 1408 compounds that were selected to be representative of the compound diversity found in the GSK HTS collection, which consists of 1.7 million compounds. Typically, at GSK, an assay configured utilising a recombinant, purified protein, would be expected to identify up to 2% of the HTS compound collection as potential hit compounds. This 2% hit rate equates to 34,000 compounds, which would, typically, be first clustered and assessed by computational chemistry tools prior to selecting 20,000 or less compounds (this number is limited by process limitations) for further hit confirmation screening[454]. The Robustness set is used as an early assay development test to ensure that the test assay hit rate is not likely to exceed 2%, which would be 14 hit compounds. If the Robustness set hit rate exceeds 2% then, typically, further assay development experiments are initiated, such as buffer and reagent optimisation. Indeed, published studies have shown that the confirmed hit rate, defined as compounds that inhibit the target of interest in at least two orthogonal assays (typically a biophysical and cellular approach) from HTS projects is much lower than 1%, as low as 0.01% to 0.14%[455]. Screening the GSK Robustness set in the GlmU acetyltransferase activity DTP absorbance assay yielded a 1.5% hit rate, which was 21 compounds. These hit rate data confirmed that the GlmU acetyltransferase activity assay was within the GSK hit identification criteria expected for a purified, recombinant protein.

The GSK Nuisance set is an annotated collection of 1027 compounds comprised of a diverse range of known assay interferences, which was compiled in response to increasing knowledge of PAINS. Screening of the GlmU acetyltransferase activity DTP absorbance assay against the Nuisance set yielded a hit rate of 22.1%, which

is high when compared to the GSK recommendations. Typically, for an assay utilising a purified, recombinant protein, a Nuisance set hit rate of between 7 to 15% would be expected, based on GSK data since the inception of this test in 2010[291]. Generally, further assay development, to reduce the hit rate, or use of an alternative screening assay would be considered before further small molecule screening would be attempted. However, the PAINS classifications of several the Nuisance set hits can be explained, and would be expected, using a thiol detection methodology to monitor the formation of CoA-SH. During a single concentration, diversity screen, using the DTP absorbance assay, hits like the PAINS identified in the Nuisance screen, would simply be excluded, or followed up with caution using a non-thiol detection methodology. Indeed, a recent publication documents the issues of utilising a thiol detection assay methodology to monitor the histone acetyltransferase, Rtt109[456]. The planned screening activities utilising the DTP absorbance assay are limited in scale, when compared to screening the GSK, 1.7 million compound HTS collection and the availability of alternative assays to confirm any hits, gave confidence to proceed using the current methodology. Additionally, awareness of the likely interference mechanisms that can afflict the absorbance assay will allow prioritisation of any hits identified from screening the diversity and focussed sets and exclusion of any PAINS inhibitors. The data generated with screening these training sets validated the use of the DTP absorbance assay for further hit identification studies, while also highlighting the likely PAINS that could be encountered.

The DTP absorbance assay was used to screen the GlmU acetyltransferase activity against the 227 compounds GSK TB set, which is a collection of compounds that have been shown to inhibit the growth of *M. tuberculosis* (this set was discussed in Section 1.4.2.6). The TB set was initially screened in duplicate at a single concentration of 10 µM, which led to the identification of 20 compounds, a hit rate of 9.6%. In parallel, the TB set compounds were screened, in duplicate, as concentration titrations, from a high concentration of 100 µM, which subsequently led to the identification of 39 hits, a hit rate of 17.2%. Reassuringly, a plot of the data generated in the single concentration screen against the IC_{50} values, determined from the dose response test occasions correlated, confirming that the same hit compounds were identified on four test occasions (Figure 101 and Table 25). Further testing of the GlmU acetyltransferase activity TB set hits from fresh preparations of small molecules confirmed the most potent inhibitors and highlighted the value of

using compounds from alternative sources. A complimentary approach to retesting compounds from different sources, is to carry out a sub-structure search based on the hit (using Tanimoto similarity[457]), and test related small molecules[458]. Testing of three small molecules related to the most potent TB set hit (by a Tanimoto similarity of between 0.93 – 0.96), confirmed that these compounds are all inhibitors of the GlmU acetyltransferase activity, with a spread of potencies. These initial data demonstrate that there is potential SAR related to this series of compounds that could be utilised to optimise a lead compound.

To confirm that the hit compounds identified by the primary screening approach, it is pertinent that, in addition to retesting compounds from an alternative source (or re-prepared stocks), the potential inhibitors are tested in at least one alternative assay. Importantly, it is key to ensure that the confirmation assays utilise an alternative methodology and, ideally, monitor either direct protein interaction of the test compound or the levels of another reactant. The approaches employed to monitor direct protein interaction of the test compound are biophysical assays, which are typically either SPR, ligand-observed or protein-observed NMR, MST or DSF[300]. The most potent GlmU acetyltransferase activity hit compound, and related small molecules, were tested in two alternative assays, which were RF-MS, utilised to monitor the concentrations of GlcN-1P and GlcNAc-1P, and DSF, employed to observe the modulation of GlmU T_m. The secondary assays confirmed that the hits were able to inhibit GlmU acetyltransferase activity, monitored by RF-MS, in addition to stabilise the DSF observed GlmU T_m.

To further explore the chemical space around the initial potent TB set hit, GSK920684A, related compounds (Tanimoto similarity > 0.8), were screened in the GlmU acetyltransferase DTP absorbance assay. This additional screening identified 25 inhibitors that had pIC50 values greater than 4.95, and further confirmation testing using MST, independently demonstrated that a subset of compounds were able to bind GlmU. MST was used to provide an alternative biophysical approach to characterise compound binding, primarily to ensure that modulation of GlmU T_m determined by DSF could be confirmed by measuring a different compound mediated protein effect. This additional screening and confirmation data supplemented the GSK920684A series dataset and provided further evidence that this series has potential for further optimisation. Crucially, for this series to be considered for further studies it was necessary to demonstrate that analogues of GSK920684A were

inhibitors *in vivo* against *M. tuberculosis*. Inhibition studies were carried out using GSK920684A and three close analogues (Tanimoto similarity between 0.93 – 0.96) against *M. tuberculosis* H37Rv, which enabled the determination of MIC values for all test compounds. GSK920684A was originally identified as an inhibitor of both *M. bovis* BCG and *M. tuberculosis* H37Rv and these data demonstrate that it is possible to conservatively modify the core chemical structure, while maintaining MIC values. The GSK TB set was filtered to ensure that all 227 compounds, including GSK920684A, in the final collection had drug-like physiochemical properties making these small molecules attractive for future lead optimisation[335,336]. These *M. tuberculosis* inhibition data validate the GSK920684A and related compounds as viable starting points for inhibitor optimisation, primarily to increase potency while maintaining desirable drug-like physiochemical properties. Furthermore, the confirmation of the GSK920684A series as dual GlmU acetyltransferase activity and *M. tuberculosis* inhibitors provides further evidence for the essentiality GlmU. To facilitate further exploration of the GSK920684A series, future studies that could be considered are mode of action experiments, such as substrate competition and residence time studies, GlmU structural biology X-ray co-crystallography with exemplar inhibitors, as well as metabolomic experiments to confirm that these hit compounds inhibit UDP-GlcNAc biosynthesis *in vivo*.

A continuation of GlmU acetyltransferase activity inhibitor identification studies was to screen additional single concentration small molecule collections, including the 3,000-compound focussed *M. tuberculosis* TB box (Figure 109) and the GSK Validation set (Figure 112), which is larger, more predictive version of the Robustness set. Screening of the aforementioned sets identified several novel potential inhibitors of the GlmU acetyltransferase activity, which further validates the chemical tractability of this active site. The TB box hits were screened in duplicate in a concentration titration experiment, leading to the identification of hit compounds with IC_{50} values less than 1 μM (Figure 111). Initially, to better validate hit compounds from these small molecule collections and confirm GlmU inhibition, testing in orthogonal acetyltransferase activity and biophysical assays is required. In the absence of any additional hit compound confirmation data, these screening experiments serve to demonstrate that both the use of the DTP thiol absorbance assay (in kinetic mode) and, more generally, the GlmU acetyltransferase activity are viable and attractive propositions for chemical tool and drug discovery.

As counterpart studies and to address some of the GlmU acetyltransferase literature conjecture, screening experiments were carried out to validate small molecules from the aforementioned reports. The premise of these studies was to explore the additional chemical space identified in these publications using the rigorous screening strategies applied to the previously discussed hit identification experiments. The ELT inhibitor identification approach taken by Machutta *et al.* at GSK broadly explored the chemical tractability of several *M. tuberculosis* proteins[310]. This study identified a limited number of GlmU-interacting compounds, which when screened against the acetyltransferase activity in the DTP thiol absorbance assay were shown to be inhibitors but of restrictive SAR. The ELT study identified GlmU as poorly chemical tractable, which may be the case using this approach, however, the results of screening experiments in this project and by other groups, such as AZ and Pfizer (using orthologues of the *M. tuberculosis* enzyme), are contrary to this classification.

Assessment of analogues of *E. coli* GlmU acetyltransferase inhibitors published by AZ demonstrated that these small molecules were poor inhibitors of the *M. tuberculosis* activity, perhaps highlighting the differences in the active sites between the two orthologues. The screening of small molecules identified as *S. aureus* GlmU acetyltransferase inhibitors was, surprisingly, much more successful, given that a 2007 publication by Payne *et al.* stated that hit identification studies with this enzyme were unsuccessful[348]. The screening of 234 chemically related *S. aureus* GlmU inhibitors against the *M. tuberculosis* enzyme identified 67% of these compounds inhibited acetyltransferase activity with IC_{50} values of greater than 100 μM. Exemplars for this chemical series had previously been shown to inhibit the growth of *S. aureus* GlmU, in addition to several other bacteria, including *E. coli, S. pneumoniae* and *H. influenzae.* Further studies are ongoing to validate whether the same exemplars are *M. tuberculosis* inhibitors, confirmation of which, given the high potency of some of these compounds, would provide an additional chemical series to further investigate alongside the TB set derived GSK920684A series.

In addition to the screening of synthetically derived small molecules, an alternative knowledge-based inhibitor design approach was explored. This early study involved the derivatisation of GlcN-1P at the C-2 amine position, to develop a potent dead-end inhibitor that would trap GlmU in a non-productive ternary complex. This approach was relatively successful, given that at the inception of the study, key

mechanistic information, such as the order of substrate binding and the acetyl transfer chemical mechanism, had yet to be elucidated. This mechanistic information would have better informed the synthesis of GlcN-1P analogues as well as the configuration of acetyltransferase activity assays to assess these compounds (ensuring the correct order of substrate addition). The modification of GlcN-1P, without further optimising active site contacts, was perhaps not an ideal starting point, given the relatively high $K_{m,\ GlcN\text{-}1P}$ and weak Glc-1P IC_{50}. Furthermore, derivatisation of the GlcN-1P amine, given the proximity to both His374 and the Ac-CoA acetyl group, was not an ideal strategy and would likely lead to steric clashes, given the substrate binding order in the acetyltransferase active site.

The screening of the GlmU acetyltransferase activity to identify small molecule inhibitors was successful, with two chemical series found. The GSK920684A series, identified from screening the GSK TB set, have been confirmed as GlmU inhibitors in several orthogonal assays. Exemplars of the GSK920684A series have been shown to inhibit the growth of *M. tuberculosis*, which confirms the essentiality of the GlmU acetyltransferase activity and UDP-GlcNAc biosynthesis.

6.4 Conclusions

The results of this project presented herein define three important facets of the kinetic and chemical mechanisms of acetyl transfer by the LβH domain of GlmU (Scheme 12 and Scheme 13). First, the intersecting acetyltransferase initial velocity patterns and product inhibition studies rule out a kinetic ping-pong mechanism and confirm that the acetylation of GlcN-1P proceed by the initial formation of a ternary complex, with Ac-CoA binding to GlmU first, followed by GlcN-1P prior to catalysis. Second, the amine of GlcN-1P is likely to be in a protonated form when it binds to GlmU, and subsequently the first essential step in the acetyl transfer reaction is the deprotonation of this amine group by an active site general base. The side chain imidazole of His374 is both optimally positioned and of the right pK_a to be able to deprotonate the positively charged amine from GlcN-1P. Mutation of His374 to an Ala, demonstrated that in the absence of the imidazole side chain, acetyl transfer only occurs when the GlcN-1P amine is in the deprotonated. There are many reports of acetyltransferase active site histidine residues acting as general bases, in addition

to catalysis proceeding by ternary complex mechanisms, some of which are detailed in Table 4 and Section 1.2.3. Third, chemistry is likely to be partially rate-limiting as both pH-rate studies and solvent kinetic isotope effects identified contributions of critical substrate and enzyme groups. Proton inventory experiments identified that one proton is in flight during the acetyl transfer reaction, the identity of which is likely to be the proton abstracted by the imidazole side chain of His374 from the NH_3^+ of GlcN-1P.

Additional key results from this project were that GlmU requires the presence of divalent metal ions, which could either be Mg^{2+}, Mn^{2+} or Ca^{2+} for optimal acetyltransferase activity. The dependence of acetyltransferase activity on divalent metal ions is concentration dependent and K_{act} values have been determined. The requirement of L*β*H acyltransferases for optimal catalysis has been shown previously for both *E. coli* GlmU and DapD[104,209].

The first identification of CoA-SH product inhibition, as well as characterisation of the novel dead-end inhibitors Bu-CoA and Glc-1P, enabled the use of inhibitor pattern studies to confirm the order of the acetyltransferase activity steady-state sequential mechanism. The first substrate likely to bind to the L*β*H domain of GlmU is Ac-CoA, which is thought to elicit a conformational change of a mobile loop, that is stabilised over the acetyl thioester of Ac-CoA forming a pocket to enable binding of GlcN-1P. These observations were confirmed by structural biology studies that captured the novel X-ray crystallography structure of the Glc-1P dead-end inhibited ternary complex. After transfer of the acetyl group to GlcN-1P, GlcNAc-1P is thought to dissociate first followed by CoA-SH. Further SDM pH-rate studies supported the observations of the X-ray crystallography structural experiments, which identified Arg344, Lys362 and Tyr377 as key residues for GlcN-1P binding.

The mechanistic features of GlmU acetyltransferase activity revealed by these studies have proved to be invaluable for both synthetic and rational inhibitor identification studies. Furthermore, these findings will provide useful basis for future GlmU mechanistic and inhibitor identification studies that could be used to finalise the essentiality of GlmU and UDP-GlcNAc biosynthesis for the optimal growth of *M. tuberculosis* and subsequently the pathogenesis of TB.

The use of drug-like small molecule inhibitors of GlmU acetyltransferase activity to explore the essentiality *in vitro* in *M. tuberculosis*, is now possible due to the identification of at least two chemical series. The identification and confirmation of

the GSK920684A series as novel inhibitors of GlmU acetyltransferase activity, was followed by the determination of MIC values for exemplars of this series against *M. tuberculosis*. These data initially demonstrate that inhibitors of GlmU acetyltransferase activity leads to inhibition of *M. tuberculosis* growth, confirming that GlmU catalysed acetyl transfer to form GlcNAc-1P, is likely to be essential. Furthermore, the chemical tractability of the acetyltransferase activity of GlmU has been demonstrated by these screening experiments, which identified several diverse, novel inhibitor starting points for future lead optimisation studies.

Chapter 7. Appendix

Reference List

1 Koch, R. Vol. Die Ätiologie der Tuberkulose (Robert Koch-Institut, 2010).

2 Gortazar, C., Delahay, R. J., Mcdonald, R. A., Boadella, M., Wilson, G. J., Gavier-Widen, D. & Acevedo, P. The status of tuberculosis in European wild mammals. *Mammal Review* **42**, 193-206 (2012).

3 Smith, I. Mycobacterium tuberculosis Pathogenesis and Molecular Determinants of Virulence. *Clinical Microbiology Reviews* **16**, 463-496 (2003).

4 de Kantor, I. N., LoBue, P. A. & Thoen, C. O. Human tuberculosis caused by Mycobacterium bovis in the United States, Latin America and the Caribbean [Serialised article. Tuberculosis: a re-emerging disease in animals and humans. Number 3 in the series]. *The International Journal of Tuberculosis and Lung Disease* **14**, 1369-1373 (2010).

5 de Jong, B. C., Antonio, M. & Gagneux, S. Mycobacterium africanum--review of an important cause of human tuberculosis in West Africa. *PLoS neglected tropical diseases* **4**, e744-e744 (2010).

6 Panteix, G., Gutierrez, M. C., Boschiroli, M. L., Rouviere, M., Plaidy, A., Pressac, D., Porcheret, H., Chyderiotis, G., Ponsada, M., Van Oortegem, K., Salloum, S., Cabuzel, S., Bañuls, A. L., Van de Perre, P. & Godreuil, S. Pulmonary tuberculosis due to Mycobacterium microti: a study of six recent cases in France. *Journal of Medical Microbiology* **59**, 984-989 (2010).

7 Bañuls, A.-L., Sanou, A., Van Anh, N. T. & Godreuil, S. Mycobacterium tuberculosis: ecology and evolution of a human bacterium. *Journal of Medical Microbiology* **64**, 1261-1269 (2015).

8 Donoghue, H. D., Lee, O. Y.-C., Minnikin, D. E., Besra, G. S., Taylor, J. H. & Spigelman, M. Tuberculosis in Dr Granville's mummy: a molecular re-examination of the earliest known Egyptian mummy to be scientifically examined and given a medical diagnosis. *Proceedings of the Royal Society B: Biological Sciences* **277**, 51-56 (2010).

9 Gagneux, S. Host pathogen coevolution in human tuberculosis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 850-859 (2012).

10 Brites, D. & Gagneux, S. Co-evolution of Mycobacterium tuberculosis and Homo sapiens. *Immunological reviews* **264**, 6-24 (2015).

11 Lienhardt, C., Glaziou, P., Uplekar, M., Lönnroth, K., Getahun, H. & Raviglione, M. Global tuberculosis control: lessons learnt and future prospects. *Nature Reviews Microbiology* **10**, 407 (2012).

12 Wilson, L. G. Commentary: Medicine, population, and tuberculosis. *International Journal of Epidemiology* **34**, 521-524 (2004).

13 Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S. R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D. & Gagneux, S. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nature Genetics* **45**, 1176 (2013).

14 WHO Global Tuberculosis Report 2018. *World Health Organisation* (2018).

15 WHO The Global MDR-TB & XDR-TB Response Plan 2007-2008. *WHO* (2007).

16 WHO Gear up to end TB: Introducing the end TB Strategy. *WHO* (2015).

17 Kaufmann, S. H. E. How can immunology contribute to the control of tuberculosis? *Nature Reviews Immunology* **1**, 20 (2001).

18 Houston, A. & Macallan, D. C. Extrapulmonary tuberculosis. *Medicine* **42**, 18-22 (2014).

19 Houben, E. N. G., Nguyen, L. & Pieters, J. Interaction of pathogenic mycobacteria with the host immune system. *Current Opinion in Microbiology* **9**, 76-85 (2006).
20 Schäfer, G., Jacobs, M., Wilkinson, R. J. & Brown, G. D. Non-Opsonic Recognition of Mycobacterium tuberculosis by Phagocytes. *Journal of Innate Immunity* **1**, 231-243 (2009).
21 Kinchen, J. M. & Ravichandran, K. S. Phagosome maturation: going through the acid test. *Nature Reviews Molecular Cell Biology* **9**, 781 (2008).
22 Bussi, C. & Gutierrez, M. G. Mycobacterium tuberculosis infection of host cells in space and time. (2019).
23 Kimmey, J. M. & Stallings, C. L. Bacterial Pathogens versus Autophagy: Implications for Therapeutic Interventions. *Trends in Molecular Medicine* **22**, 1060-1076 (2016).
24 Lerner, T. R., Borel, S., Greenwood, D. J., Repnik, U., Russell, M. R. G., Herbst, S., Jones, M. L., Collinson, L. M., Griffiths, G. & Gutierrez, M. G. Mycobacterium tuberculosis replicates within necrotic human macrophages. *The Journal of Cell Biology* **216**, 583-594 (2017).
25 Pajuelo, D., Gonzalez-Juarbe, N., Tak, U., Sun, J., Orihuela, C. J. & Niederweis, M. NAD+ Depletion Triggers Macrophage Necroptosis, a Cell Death Pathway Exploited by Mycobacterium tuberculosis. *Cell Reports* **24**, 429-440 (2018).
26 Cadena, A. M., Fortune, S. M. & Flynn, J. L. Heterogeneity in tuberculosis. *Nature Reviews Immunology* **17**, 691 (2017).
27 Jackson, M. The Mycobacterial Cell Envelope—Lipids. *Cold Spring Harbor Perspectives in Medicine* (2014).
28 Jankute, M., Cox, J. A. G., Harrison, J. & Besra, G. S. Assembly of the Mycobacterial Cell Wall. *Annual Review of Microbiology* **69**, 405-423 (2015).
29 Daffé, M. The cell envelope of tubercle bacilli. *Tuberculosis* **95**, S155-S158 (2015).
30 Vincent, A. T., Nyongesa, S., Morneau, I., Reed, M. B., Tocheva, E. I. & Veyrier, F. J. The Mycobacterial Cell Envelope: A Relict From the Past or the Result of Recent Evolution? *Frontiers in microbiology* **9**, 2341-2341 (2018).
31 Srivastava, S., van Rijn, S. P., Wessels, A. M. A., Alffenaar, J.-W. C. & Gumbo, T. Susceptibility Testing of Antibiotics That Degrade Faster than the Doubling Time of Slow-Growing Mycobacteria: Ertapenem Sterilizing Effect versus Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy* **60**, 3193-3195 (2016).
32 Sezonov, G., Joseleau-Petit, D. & D'Ari, R. Escherichia coli Physiology in Luria-Bertani Broth. *Journal of Bacteriology* **189**, 8746-8749 (2007).
33 Lory, S. in *The Prokaryotes: Actinobacteria* (eds Eugene Rosenberg *et al.*) 571-575 (Springer Berlin Heidelberg, 2014).
34 Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omaïs, B., Marmiesse, M., Supply, P. & Vincent, V. Ancient Origin and Gene Mosaicism of the Progenitor of Mycobacterium tuberculosis. *PLOS Pathogens* **1**, e5 (2005).
35 Wirth, T., Hildebrand, F., Allix-Béguec, C., Wölbeling, F., Kubica, T., Kremer, K., van Soolingen, D., Rüsch-Gerdes, S., Locht, C., Brisse, S., Meyer, A., Supply, P. & Niemann, S. Origin, Spread and Demography of the Mycobacterium tuberculosis Complex. *PLOS Pathogens* **4**, e1000160 (2008).
36 Mellars, P. Going East: New Genetic and Archaeological Perspectives on the Modern Human Colonization of Eurasia. *Science* **313**, 796-800 (2006).
37 Masson, M., Bereczki, Z., Molnár, E., Donoghue, H. D., Minnikin, D. E., Lee, O. Y. C., Wu, H. H. T., Besra, G. S., Bull, I. D. & Pálfi, G. 7000 year-old

tuberculosis cases from Hungary – Osteological and biomolecular evidence. *Tuberculosis* **95**, S13-S17 (2015).

38 Barberis, I., Bragazzi, N. L., Galluzzo, L. & Martini, M. The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. *Journal of preventive medicine and hygiene* **58**, E9-E12 (2017).

39 Doetsch, R. N. Benjamin Marten and his "New Theory of Consumptions". *Microbiological reviews* **42**, 521-528 (1978).

40 Cambau, E. & Drancourt, M. Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. *Clinical Microbiology and Infection* **20**, 196-201 (2014).

41 Blevins, S. M. & Bronze, M. S. Robert Koch and the 'golden age' of bacteriology. *International Journal of Infectious Diseases* **14**, e744-e751 (2010).

42 Pavelka Jr., M. S., Mahapatra, S. & Crick, D. C. Genetics of Peptidoglycan Biosynthesis. *Microbiology Spectrum* **2** (2014).

43 Schleifer, K. H. & Kandler, O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological reviews* **36**, 407-477 (1972).

44 Alderwick, L. J., Harrison, J., Lloyd, G. S. & Birch, H. L. The Mycobacterial Cell Wall--Peptidoglycan and Arabinogalactan. *Cold Spring Harbor perspectives in medicine* **5**, a021113-a021113.

45 Kieser, K. J. & Rubin, E. J. How sisters grow apart: mycobacterial growth and division. *Nature Reviews Microbiology* **12**, 550 (2014).

46 Chiaradia, L., Lefebvre, C., Parra, J., Marcoux, J., Burlet-Schiltz, O., Etienne, G., Tropis, M. & Daffé, M. Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane. *Scientific Reports* **7**, 12807 (2017).

47 Zumla, A., Nahid, P. & Cole, S. T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nature Reviews Drug Discovery* **12**, 388 (2013).

48 Onyebujoh, P., Zumla, A., Ribeiro, I., Rustomjee, R., Mwaba, P., Gomes, M. & Grange, J. Treatment of tuberculosis: present status and future prospects. *Bulletin of the World Health Organization* **83**, 857-865 (2005).

49 Tuberculosis Coalition for Technical Assistance. . *International Standards for Tuberculosis Care (Tuberculosis Coalition for Technical Assistance, 2009)* **2nd edn** (2009).

50 E. Barry, C. Lessons from Seven Decades of Antituberculosis Drug Discovery. *Current Topics in Medicinal Chemistry* **11**, 1216-1225 (2011).

51 Gengiah, T. N., Gray, A. L., Naidoo, K. & Karim, Q. A. Initiating antiretrovirals during tuberculosis treatment: a drug safety review. *Expert opinion on drug safety* **10**, 559-574 (2011).

52 Thummel, K. E. & Wilkinson, G. R. In vitro and in vivo drug interactions involving human CYP3A. *Annual Review of Pharmacology and Toxicology* **38**, 389-430 (1998).

53 Finch, C. K., Chrisman, C. R., Baciewicz, A. M. & Self, T. H. Rifampin and Rifabutin Drug Interactions: An Update. *Archives of Internal Medicine* **162**, 985-992 (2002).

54 Falzon, D., Jaramillo, E., Schünemann, H. J., Arentz, M., Bauer, M., Bayona, J., Blanc, L., Caminero, J. A., Daley, C. L., Duncombe, C., Fitzpatrick, C., Gebhard, A., Getahun, H., Henkens, M., Holtz, T. H., Keravec, J., Keshavjee, S., Khan, A. J., Kulier, R., Leimane, V., Lienhardt, C., Lu, C., Mariandyshev, A., Migliori, G. B., Mirzayev, F., Mitnick, C. D., Nunn, P., Nwagboniwe, G., Oxlade, O., Palmero, D., Pavlinac, P., Quelapio, I., Raviglione, M. C., Rich, M. L., Royce, S., Rüsch-Gerdes, S., Salakaia, A., Sarin, R., Sculier, D., Varaine, F., Vitoria, M., Walson, J. L., Wares, F., Weyer, K., White, R. A. & Zignol, M. WHO

guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update. *European Respiratory Journal*, erj00736-02011 (2011).

55 WHO. WHO Rapid Communication: Key changes to treatment of multidrug- and rifampicin-resistant tuberculosis (MDR/RR-TB) (2018).

56 Reeves, A. Z., Campbell, P. J., Willby, M. J. & Posey, J. E. Disparities in Capreomycin Resistance Levels Associated with the A1401G Mutation in Clinical Isolates of Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy* **59**, 444-449 (2015).

57 Guglielmetti, L., Barkane, L., Le Du, D., Marigot-Outtandy, D., Robert, J., Veziris, N., Yazdanpanah, Y., Kuksa, L., Caumes, E. & Fréchet-Jachym, M. Safety and efficacy of exposure to bedaquiline-delamanid in MDR-TB: a case series from France and Latvia. *European Respiratory Journal*, 1702550 (2018).

58 Chee, C. B. E., KhinMar, K.-W., Sng, L.-H., Jureen, R., Cutter, J., Lee, V. J. M. & Wang, Y.-T. The shorter multidrug-resistant tuberculosis treatment regimen in Singapore: are patients from South-East Asia eligible? *European Respiratory Journal* **50**, 1700753 (2017).

59 Caminero, J. A., Piubello, A., Scardigli, A. & Migliori, Giovanni B. Proposal for a standardised treatment regimen to manage pre- and extensively drug-resistant tuberculosis cases. *European Respiratory Journal* **50**, 1700648 (2017).

60 Parida, S. K., Axelsson-Robertson, R., Rao, M. V., Singh, N., Master, I., Lutckii, A., Keshavjee, S., Andersson, J., Zumla, A. & Maeurer, M. Totally drug-resistant tuberculosis and adjunct therapies. *Journal of Internal Medicine* **277**, 388-405 (2015).

61 Migliori, G. B., Sotgiu, G., Gandhi, N. R., Falzon, D., DeRiemer, K., Centis, R., Hollm-Delgado, M.-G., Palmero, D., Pérez-Guzmán, C., Vargas, M. H., D'Ambrosio, L., Spanevello, A., Bauer, M., Chan, E. D., Schaaf, H. S., Keshavjee, S., Holtz, T. H. & Menzies, D. Drug resistance beyond extensively drug-resistant tuberculosis: individual patient data meta-analysis. *European Respiratory Journal* **42**, 169-179 (2013).

62 Laurence, Y. V., Griffiths, U. K. & Vassall, A. Costs to Health Services and the Patient of Treating Tuberculosis: A Systematic Literature Review. *PharmacoEconomics* **33**, 939-955 (2015).

63 Raviglione, M., Uplekar, M., Weil, D. & Kasaeva, T. Tuberculosis makes it onto the international political agenda for health finally. *The Lancet Global Health* **6**, e20-e21 (2018).

64 TAG 2016 Report on Tuberculosis Research Funding Trends, 2005–2015: No Time to Lose. *TAG* (2016).

65 Cohen, J. Approval of Novel TB Drug Celebrated—With Restraint. *Science* **339**, 130-130 (2013).

66 Gler, M. T., Skripconoka, V., Sanchez-Garavito, E., Xiao, H., Cabrera-Rivero, J. L., Vargas-Vasquez, D. E., Gao, M., Awad, M., Park, S.-K., Shim, T. S., Suh, G. Y., Danilovits, M., Ogata, H., Kurve, A., Chang, J., Suzuki, K., Tupasi, T., Koh, W.-J., Seaworth, B., Geiter, L. J. & Wells, C. D. Delamanid for Multidrug-Resistant Pulmonary Tuberculosis. *New England Journal of Medicine* **366**, 2151-2160 (2012).

67 Lohrasbi, V., Talebi, M., Bialvaei, A. Z., Fattorini, L., Drancourt, M., Heidary, M. & Darban-Sarokhalil, D. Trends in the discovery of new drugs for Mycobacterium tuberculosis therapy with a glance at resistance. *Tuberculosis* **109**, 17-27 (2018).

68 Ahsan, M. J. Recent advances in the development of vaccines for tuberculosis. *Therapeutic advances in vaccines* **3**, 66-75 (2015).

69 McNerney, R., Cunningham, J., Hepple, P. & Zumla, A. New tuberculosis diagnostics and rollout. *International Journal of Infectious Diseases* **32**, 81-86 (2015).
70 Martínez Cuesta, S., Rahman, S. A., Furnham, N. & Thornton, J. M. The Classification and Evolution of Enzyme Function. *Biophysical journal* **109**, 1082-1086 (2015).
71 Roth, S. Y., Denu, J. M. & Allis, C. D. Histone Acetyltransferases. *Annual Review of Biochemistry* **70**, 81-120 (2001).
72 Chen, W., Biswas, T., Porter, V. R., Tsodikov, O. V. & Garneau-Tsodikova, S. Unusual regioversatility of acetyltransferase Eis, a cause of drug resistance in XDR-TB. *Proceedings of the National Academy of Sciences* **108**, 9804-9808 (2011).
73 Vetting, M. W., S. de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L. & Blanchard, J. S. Structure and functions of the GNAT superfamily of acetyltransferases. *Archives of Biochemistry and Biophysics* **433**, 212-226 (2005).
74 Berndsen, C. E., Albaugh, B. N., Tan, S. & Denu, J. M. Catalytic Mechanism of a MYST Family Histone Acetyltransferase. *Biochemistry* **46**, 623-629 (2007).
75 Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E. & Noble, M. E. M. Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nature Structural Biology* **7**, 560 (2000).
76 Sikora, A. L., Frankel, B. A. & Blanchard, J. S. Kinetic and Chemical Mechanism of Arylamine N-Acetyltransferase from Mycobacterium tuberculosis. *Biochemistry* **47**, 10781-10789 (2008).
77 Bogicevic, B., Berthoud, H., Portmann, R., Bavan, T., Irmler, S. & Meile, L. Cysteine biosynthesis in Lactobacillus casei: identification and characterization of a serine acetyltransferase. *FEMS Microbiology Letters* **363** (2016).
78 Pye, V. E., Tingey, A. P., Robson, R. L. & Moody, P. C. E. The Structure and Mechanism of Serine Acetyltransferase from Escherichia coli. *Journal of Biological Chemistry* **279**, 40729-40736 (2004).
79 Wang, X.-G., Olsen, L. R. & Roderick, S. L. Structure of the lac Operon Galactoside Acetyltransferase. *Structure* **10**, 581-588 (2002).
80 Nguyen, L., Kozlov, G. & Gehring, K. Structure of Escherichia coli tetrahydrodipicolinate N-succinyltransferase reveals the role of a conserved C-terminal helix in cooperative substrate binding. *FEBS Letters* **582**, 623-626 (2008).
81 Lo Leggio, L., Dal Degan, F., Poulsen, P., Andersen, S. M. & Larsen, S. The Structure and Specificity of Escherichia coli Maltose Acetyltransferase Give New Insight into the LacA Family of Acyltransferases. *Biochemistry* **42**, 5225-5235 (2003).
82 Beaman, T. W., Sugantino, M. & Roderick, S. L. Structure of the Hexapeptide Xenobiotic Acetyltransferase from Pseudomonas aeruginosa. *Biochemistry* **37**, 6689-6696 (1998).
83 Williams, A. H., Immormino, R. M., Gewirth, D. T. & Raetz, C. R. H. Structure of UDP-N-acetylglucosamine acyltransferase with a bound antibacterial pentadecapeptide. *Proceedings of the National Academy of Sciences* **103**, 10877-10882 (2006).
84 Buetow, L., Smith, T. K., Dawson, A., Fyffe, S. & Hunter, W. N. Structure and reactivity of LpxD, the N-acyltransferase of lipid A biosynthesis. *Proceedings of the National Academy of Sciences* **104**, 4321-4326 (2007).
85 Alber, B. E. & Ferry, J. G. A carbonic anhydrase from the archaeon Methanosarcina thermophila. *Proceedings of the National Academy of Sciences* **91**, 6909-6913 (1994).

86 Raetz, C. R. H. & Roderick, S. L. A Left-Handed Parallel β Helix in the Structure of UDP-N-Acetylglucosamine Acyltransferase. *Science* **270**, 997-1000 (1995).
87 Dicker, I. B. & Seetharam, S. What is known about the structure and function of the Escherichia coli protein FirA? *Molecular Microbiology* **6**, 817-823 (1992).
88 Vaara, M. Eight bacterial proteins, including UDP-N-acetylglucosamine acyltransferase (LpxA) and three other transferases of Escherichia coli, consist of a six-residue periodicity theme. *FEMS Microbiology Letters* **97**, 249-254 (1992).
89 Beaman, T. W., Binder, D. A., Blanchard, J. S. & Roderick, S. L. Three-Dimensional Structure of Tetrahydrodipicolinate N-Succinyltransferase. *Biochemistry* **36**, 489-494 (1997).
90 Olsen, L. R., Huang, B., Vetting, M. W. & Roderick, S. L. Structure of Serine Acetyltransferase in Complexes with CoA and Its Cysteine Feedback Inhibitor. *Biochemistry* **43**, 6013-6019 (2004).
91 Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B. & Jez, J. M. Structure of Soybean Serine Acetyltransferase and Formation of the Cysteine Regulatory Complex as a Molecular Chaperone. *Journal of Biological Chemistry* **288**, 36463-36472 (2013).
92 Luo, H.-B., Knapik, A. A., Petkowski, J. J., Demas, M., Shumilin, I. A., Zheng, H., Chruszcz, M. & Minor, W. Biophysical analysis of the putative acetyltransferase SACOL2570 from methicillin-resistant Staphylococcus aureus. *Journal of structural and functional genomics* **14**, 97-108 (2013).
93 Sugantino, M. & Roderick, S. L. Crystal Structure of Vat(D): An Acetyltransferase That Inactivates Streptogramin Group A Antibiotics. *Biochemistry* **41**, 2209-2216 (2002).
94 Lee, B. I. & Suh, S. W. Crystal structure of UDP-N-acetylglucosamine acyltransferase fromHelicobacter pylori. *Proteins: Structure, Function, and Bioinformatics* **53**, 772-774 (2003).
95 Robins, L. I., Williams, A. H. & Raetz, C. R. H. Structural Basis for the Sugar Nucleotide and Acyl-Chain Selectivity of Leptospira interrogans LpxA. *Biochemistry* **48**, 6191-6201 (2009).
96 Badger, J., Chie-Leon, B., Logan, C., Sridhar, V., Sankaran, B., Zwart, P. H. & Nienaber, V. Structure determination of LpxA from the lipopolysaccharide-synthesis pathway of Acinetobacter baumannii. *Acta crystallographica. Section F, Structural biology and crystallization communications* **68**, 1477-1481 (2012).
97 Joo, S. H., Chung, H. S., Raetz, C. R. H. & Garrett, T. A. Activity and Crystal Structure of Arabidopsis thalianaUDP-N-Acetylglucosamine Acyltransferase. *Biochemistry* **51**, 4322-4330 (2012).
98 Baugh, L., Gallagher, L. A., Patrapuvich, R., Clifton, M. C., Gardberg, A. S., Edwards, T. E., Armour, B., Begley, D. W., Dieterich, S. H., Dranow, D. M., Abendroth, J., Fairman, J. W., Fox, D., III, Staker, B. L., Phan, I., Gillespie, A., Choi, R., Nakazawa-Hewitt, S., Nguyen, M. T., Napuli, A., Barrett, L., Buchko, G. W., Stacy, R., Myler, P. J., Stewart, L. J., Manoil, C. & Van Voorhis, W. C. Combining Functional and Structural Genomics to Sample the Essential Burkholderia Structome. *PLOS ONE* **8**, e53851 (2013).
99 Joo, S. H. & Chung, H. S. Crystal structure and activity of Francisella novicida UDP-N-acetylglucosamine acyltransferase. *Biochemical and Biophysical Research Communications* **478**, 1223-1229 (2016).
100 Smith, E. W., Zhang, X., Behzadi, C., Andrews, L. D., Cohen, F. & Chen, Y. Structures of Pseudomonas aeruginosa LpxA Reveal the Basis for Its Substrate Selectivity. *Biochemistry* **54**, 5937-5948 (2015).
101 Ngo, A., Fong, K. T., Cox, D. L., Chen, X. & Fisher, A. J. Structures of Bacteroides fragilis uridine 5'-diphosphate-N-acetylglucosamine (UDP-GlcNAc)

acyltransferase (BfLpxA). *Acta Crystallographica Section D* **71**, 1068-1076 (2015).

102 Bartling, C. M. & Raetz, C. R. H. Crystal Structure and Acyl Chain Selectivity of Escherichia coli LpxD, the N-Acyltransferase of Lipid A Biosynthesis. *Biochemistry* **48**, 8672-8683 (2009).

103 Badger, J., Chie-Leon, B., Logan, C., Sridhar, V., Sankaran, B., Zwart, P. H. & Nienaber, V. The structure of LpxD from Pseudomonas aeruginosa at 1.3 Å resolution. *Acta crystallographica. Section F, Structural biology and crystallization communications* **67**, 749-752 (2011).

104 Schuldt, L., Weyand, S., Kefala, G. & Weiss, M. S. The Three-dimensional Structure of a Mycobacterial DapD Provides Insights into DapD Diversity and Reveals Unexpected Particulars about the Enzymatic Mechanism. *Journal of Molecular Biology* **389**, 863-879 (2009).

105 Schnell, R., Oehlmann, W., Sandalova, T., Braun, Y., Huck, C., Maringer, M., Singh, M. & Schneider, G. Tetrahydrodipicolinate N-Succinyltransferase and Dihydrodipicolinate Synthase from Pseudomonas aeruginosa: Structure Analysis and Gene Deletion. *PLOS ONE* **7**, e31133 (2012).

106 Sagong, H.-Y. & Kim, K.-J. Crystal Structure and Biochemical Characterization of Tetrahydrodipicolinate N-Succinyltransferase from Corynebacterium glutamicum. *Journal of Agricultural and Food Chemistry* **63**, 10641-10646 (2015).

107 Thoden, J. B., Cook, P. D., Schäffer, C., Messner, P. & Holden, H. M. Structural and Functional Studies of QdtC: An N-Acetyltransferase Required for the Biosynthesis of dTDP-3-Acetamido-3,6-dideoxy-α-d-glucose. *Biochemistry* **48**, 2699-2709 (2009).

108 Thoden, J. B. & Holden, H. M. Molecular Structure of WlbB, a Bacterial N-Acetyltransferase Involved in the Biosynthesis of 2,3-Diacetamido-2,3-dideoxy-d-mannuronic Acid. *Biochemistry* **49**, 4644-4653 (2010).

109 Chantigian, D. P., Thoden, J. B. & Holden, H. M. Structural and Biochemical Characterization of a Bifunctional Ketoisomerase/N-Acetyltransferase from Shewanella denitrificans. *Biochemistry* **52**, 8374-8385 (2013).

110 Olivier, N. B. & Imperiali, B. Crystal Structure and Catalytic Mechanism of PglD from Campylobacter jejuni. *Journal of Biological Chemistry* **283**, 27937-27946 (2008).

111 Morrison, M. J. & Imperiali, B. Biochemical Analysis and Structure Determination of Bacterial Acetyltransferases Responsible for the Biosynthesis of UDP-N,N′-Diacetylbacillosamine. *Journal of Biological Chemistry* **288**, 32248-32260 (2013).

112 Stogios, P. J., Kuhn, M. L., Evdokimova, E., Courvalin, P., Anderson, W. F. & Savchenko, A. Potential for reduction of streptogramin A resistance revealed by structural analysis of acetyltransferase VatA. *Antimicrobial agents and chemotherapy* **58**, 7083-7092 (2014).

113 Thoden, J. B., Reinhardt, L. A., Cook, P. D., Menden, P., Cleland, W. W. & Holden, H. M. The Catalytic Mechanism of Perosamine N-Acetyltransferase Revealed by High Resolution X-ray Crystallographic Studies and Kinetic Analyses(). *Biochemistry* **51**, 3433-3444 (2012).

114 Honda, Y., Nakano, S., Ito, S., Dadashipour, M., Zhang, Z. & Kawarabayasi, Y. Improvement of ST0452 N-Acetylglucosamine-1-Phosphate Uridyltransferase Activity by the Cooperative Effect of Two Single Mutations Identified through Structure-Based Protein Engineering. *Applied and Environmental Microbiology* **84**, e02213-02218 (2018).

115 Kubiak, R. L. & Holden, H. M. Structural Studies of AntD: An N-Acyltransferase Involved in the Biosynthesis of d-Anthrose. *Biochemistry* **51**, 867-878 (2012).
116 Johnson, C. M., Huang, B., Roderick, S. L. & Cook, P. F. Kinetic mechanism of the serine acetyltransferase from Haemophilus influenzae. *Archives of Biochemistry and Biophysics* **429**, 115-122 (2004).
117 Ferry, J. G. The γ class of carbonic anhydrases. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1804**, 374-381 (2010).
118 Supuran, C. T. & Capasso, C. An Overview of the Bacterial Carbonic Anhydrases. *Metabolites* **7**, 56 (2017).
119 Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. A left-handed β-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon Methanosarcina thermophila. *EMBO Journal* **15**, 2323-2330 (1996).
120 Peña, K. L., Castel, S. E., de Araujo, C., Espie, G. S. & Kimber, M. S. Structural basis of the oxidative activation of the carboxysomal γ-carbonic anhydrase, CcmM. *Proceedings of the National Academy of Sciences* **107**, 2455-2460 (2010).
121 Park, H.-M., Park, J.-H., Choi, J.-W., Lee, J., Kim, B. Y., Jung, C.-H. & Kim, J.-S. Structures of the [gamma]-class carbonic anhydrase homologue YrdA suggest a possible allosteric switch. *Acta Crystallographica Section D* **68**, 920-926 (2012).
122 Herrou, J. & Crosson, S. Molecular Structure of the Brucella abortus Metalloprotein RicA, a Rab2-Binding Virulence Effector. *Biochemistry* **52**, 9020-9028 (2013).
123 Jeyakanthan, J., Rangarajan, S., Mridula, P., Kanaujia, S. P., Shiro, Y., Kuramitsu, S., Yokoyama, S. & Sekar, K. Observation of a calcium-binding site in the [gamma]-class carbonic anhydrase from Pyrococcus horikoshii. *Acta Crystallographica Section D* **64**, 1012-1019 (2008).
124 Leu, L. S. & Cook, P. F. Kinetic Mechanism of Serine Transacetylase from Salmonella typhimurium. *Biochemistry* **33**, 2667-2671 (1994).
125 Johnson, C. M., Roderick, S. L. & Cook, P. F. The serine acetyltransferase reaction: acetyl transfer from an acylpantothenyl donor to an alcohol. *Archives of Biochemistry and Biophysics* **433**, 85-95 (2005).
126 Lewendon, A., Ellis, J. & Shaw, W. V. Structural and Mechanistic Studies of Galactoside Acetyltransferase, the Escherichia coli LacA Gene Product. *Journal of Biological Chemistry* **270**, 26326-26331 (1995).
127 Johnson, C. M., Huang, B., Roderick, S. L. & Cook, P. F. Chemical Mechanism of the Serine Acetyltransferase from Haemophilus influenzae. *Biochemistry* **43**, 15534-15539 (2004).
128 Guan, R., Roderick, S. L., Huang, B. & Cook, P. F. Roles of Histidines 154 and 189 and Aspartate 139 in the Active Site of Serine Acetyltransferase from Haemophilus influenzae. *Biochemistry* **47**, 6322-6328 (2008).
129 Rangarajan, E. S., Ruane, K. M., Sulea, T., Watson, D. C., Proteau, A., Leclerc, S., Cygler, M., Matte, A. & Young, N. M. Structure and Active Site Residues of PglD, an N-Acetyltransferase from the Bacillosamine Synthetic Pathway Required for N-Glycan Synthesis in Campylobacter jejuni. *Biochemistry* **47**, 1827-1836 (2008).
130 Wu, D., Govindasamy, L., Lian, W., Gu, Y., Kukar, T., Agbandje-McKenna, M. & McKenna, R. Structure of Human Carnitine Acetyltransferase: molecular basis for fatty acyl transfer. *Journal of Biological Chemistry* **278**, 13159-13165 (2003).
131 Ramsay, R. R. & Naismith, J. H. A snapshot of carnitine acetyltransferase. *Trends in Biochemical Sciences* **28**, 343-346 (2003).

132 Leslie, A. G., Moody, P. C. & Shaw, W. V. Structure of chloramphenicol acetyltransferase at 1.75-A resolution. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 4133-4137 (1988).
133 W V Shaw, a. & Leslie, A. G. W. Chloramphenicol Acetyltransferase. *Annual Review of Biophysics and Biophysical Chemistry* **20**, 363-386 (1991).
134 Kim, A.-R., Rylett, R. J. & Shilton, B. H. Substrate Binding and Catalytic Mechanism of Human Choline Acetyltransferase. *Biochemistry* **45**, 14621-14631 (2006).
135 Tanner, K. G., Langer, M. R. & Denu, J. M. Kinetic Mechanism of Human Histone Acetyltransferase P/CAF. *Biochemistry* **39**, 11961-11969 (2000).
136 Clements, A., Rojas, J. R., Trievel, R. C., Wang, L., Berger, S. L. & Marmorstein, R. Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A. *The EMBO journal* **18**, 3521-3532 (1999).
137 Hickman, A. B., Namboodiri, M. A. A., Klein, D. C. & Dyda, F. The Structural Basis of Ordered Substrate Binding by Serotonin N-Acetyltransferase: Enzyme Complex at 1.8 Å Resolution with a Bisubstrate Analog. *Cell* **97**, 361-369 (1999).
138 Mirza, I. A., Nazi, I., Korczynska, M., Wright, G. D. & Berghuis, A. M. Crystal Structure of Homoserine Transacetylase from Haemophilus influenzae Reveals a New Family of α/β-Hydrolases. *Biochemistry* **44**, 15768-15773 (2005).
139 Born, T. L., Franklin, M. & Blanchard, J. S. Enzyme-Catalyzed Acylation of Homoserine: Mechanistic Characterization of the Haemophilus influenzae met2-Encoded Homoserine Transacetylase. *Biochemistry* **39**, 8556-8564 (2000).
140 Born, T. L. & Blanchard, J. S. Enzyme-Catalyzed Acylation of Homoserine: Mechanistic Characterization of the Escherichia coli metA-Encoded Homoserine Transsuccinylase. *Biochemistry* **38**, 14416-14423 (1999).
141 Zubieta, C., Krishna, S. S., McMullan, D., Miller, M. D., Abdubek, P., Agarwalla, S., Ambing, E., Astakhova, T., Axelrod, H. L., Carlton, D., Chiu, H.-J., Clayton, T., Deller, M., DiDonato, M., Duan, L., Elsliger, M.-A., Grzechnik, S. K., Hale, J., Hampton, E., Han, G. W., Haugen, J., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Koesema, E., Kumar, A., Marciano, D., Morse, A. T., Nigoghossian, E., Oommachen, S., Reyes, R., Rife, C. L., Bedem, H. v. d., Weekes, D., White, A., Xu, Q., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A. & Wilson, I. A. Crystal structure of homoserine O-succinyltransferase from Bacillus cereus at 2.4 Å resolution. *Proteins: Structure, Function, and Bioinformatics* **68**, 999-1005 (2007).
142 Abuhammad, A., Lowe, E. D., McDonough, M. A., Shaw Stewart, P. D., Kolek, S. A., Sim, E. & Garman, E. F. Structure of arylamine N-acetyltransferase from Mycobacterium tuberculosis determined by cross-seeding with the homologous protein from M. marinum: triumph over adversity. *Acta Crystallographica Section D* **69**, 1433-1446 (2013).
143 Walters, K., Boukouvala, S. & Sim, E. Arylamine N-acetyltransferases: From Structure to Function AU. *Drug Metabolism Reviews* **40**, 479-510 (2008).
144 Manthei, K. A., Ahn, J., Glukhova, A., Yuan, W., Larkin, C., Manett, T. D., Chang, L., Shayman, J. A., Axley, M. J., Schwendeman, A. & Tesmer, J. J. G. A retractable lid in lecithin:cholesterol acyltransferase provides a structural mechanism for activation by apolipoprotein A-I. *Journal of Biological Chemistry* **292**, 20313-20327 (2017).
145 Peelman, F., Vinaimont, N., Vanloo, B., Labeur, C., Rosseneu, M., Verhee, A., Verschelde, J.-L., Vanderckove, J., Tavernier, J., Seguret-Mace, S., Duverger, N. & Hutchinson, G. A proposed architecture for lecithin cholesterol acyl

transferase (LCAT): Identification of the catalytic triad and molecular modeling. *Protein Science* **7**, 587-599 (1998).
146 Jogl, G. & Tong, L. Crystal Structure of Carnitine Acetyltransferase and Implications for the Catalytic Mechanism and Fatty Acid Transport. *Cell* **112**, 113-122 (2003).
147 Lewendon, A., Murray, I. A., Kleanthous, C., Cullis, P. M. & Shaw, W. V. Substitutions in the active site of chloramphenicol acetyltransferase: role of a conserved aspartate. *Biochemistry* **27**, 7385-7390 (1988).
148 Kresge, N., Simoni, R. D. & Hill, R. L. Selman Waksman: the Father of Antibiotics. *Journal of Biological Chemistry* **279**, e7 (2004).
149 Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiology and molecular biology reviews : MMBR* **74**, 417-433 (2010).
150 Harvey, A. L. Natural products in drug discovery. *Drug Discovery Today* **13**, 894-901 (2008).
151 Wright, G. D. Molecular mechanisms of antibiotic resistance. *Chem Commun (Camb)* **47**, 4055-4061 (2011).
152 Genilloud, O. Actinomycetes: still a source of novel antibiotics. *Natural Product Reports* **34**, 1203-1232 (2017).
153 Favrot, L., Blanchard, J. S. & Vergnolle, O. Bacterial GCN5-Related N-Acetyltransferases: From Resistance to Regulation. *Biochemistry* **55**, 989-1002 (2016).
154 Park, C. H., Robicsek, A., Jacoby, G. A., Sahm, D. & Hooper, D. C. Prevalence in the United States of aac(6′)-Ib-cr Encoding a Ciprofloxacin-Modifying Enzyme. *Antimicrobial Agents and Chemotherapy* **50**, 3953-3955 (2006).
155 Ehrlich, J., Bartz, Q. R., Smith, R. M., Joslyn, D. A. & Burkholder, P. R. Chloromycetin, a New Antibiotic From a Soil Actinomycete. *Science* **106**, 417-417 (1947).
156 Yunis, A. A. Chloramphenicol: Relation of Structure to Activity and Toxicity. *Annual Review of Pharmacology and Toxicology* **28**, 83-100 (1988).
157 Dunkle, J. A., Xiong, L., Mankin, A. S. & Cate, J. H. D. Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proceedings of the National Academy of Sciences* **107**, 17152-17157 (2010).
158 Cocito, C. Antibiotics of the virginiamycin family, inhibitors which contain synergistic components. *Microbiological Reviews* **43**, 145-192 (1979).
159 Thal, L. A. & Zervos, M. J. Occurrence and epidemiology of resistance to virginiamycin and streptogramins. *Journal of Antimicrobial Chemotherapy* **43**, 171-176 (1999).
160 Johnston, N. J., Mukhtar, T. A. & Wright, G. D. Streptogramin Antibiotics: Mode of Action and Resistance. *Current Drug Targets* **3**, 335-344 (2002).
161 Tennigkeit, J. & Matzura, H. Nucleotide sequence analysis of a chloramphenicol-resistance determinant from Agrobacterium tumefaciens and identification of its gene product. *Gene* **98**, 113-116 (1991).
162 Rende-Fournier, R., Leclercq, R., Galimand, M., Duval, J. & Courvalin, P. Identification of the satA gene encoding a streptogramin A acetyltransferase in Enterococcus faecium BM4145. *Antimicrobial Agents and Chemotherapy* **37**, 2119-2125 (1993).
163 Tian, Y., Beaman, T. W. & Roderick, S. L. Purification and crystallization of Pseudomonas aeruginosa chloramphenicol acetyltransferase. *Proteins: Structure, Function, and Bioinformatics* **28**, 298-300 (1997).
164 Silhavy, T. J., Kahne, D. & Walker, S. The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology* **2** (2010).

165 Vollmer, W., Blanot, D. & De Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews* **32**, 149-167 (2008).
166 Angala, S. K., Belardinelli, J. M., Huc-Claustre, E., Wheat, W. H. & Jackson, M. The cell envelope glycoconjugates of Mycobacterium tuberculosis. *Critical Reviews in Biochemistry and Molecular Biology* **49**, 361-399 (2014).
167 Hrast, M., Jukič, M., Patin, D., Tod, J., Dowson, C. G., Roper, D. I., Barreteau, H. & Gobec, S. In silico identification, synthesis and biological evaluation of novel tetrazole inhibitors of MurB. *Chemical Biology & Drug Design* **91**, 1101-1112 (2018).
168 Boniface, A., Barreteau, H., Blanot, D., Kovač, A., Sova, M. & Gobec, S. Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews* **32**, 168-207 (2008).
169 Bugg, T. D. H. & Walsh, C. T. Intracellular steps of bacterial cell wall peptidoglycan biosynthesis: enzymology, antibiotics, and antibiotic resistance. *Natural Product Reports* **9**, 199-215 (1992).
170 Durand, P., Golinelli-Pimpaneau, B., Mouilleron, S., Badet, B. & Badet-Denisot, M.-A. Highlights of glucosamine-6P synthase catalysis. *Archives of Biochemistry and Biophysics* **474**, 302-317 (2008).
171 Mouilleron, S., Badet-Denisot, M.-A. & Golinelli-Pimpaneau, B. Glutamine Binding Opens the Ammonia Channel and Activates Glucosamine-6P Synthase. *Journal of Biological Chemistry* **281**, 4404-4412 (2006).
172 Badet, B., Vermoote, P. & Le Goffic, F. Glucosamine synthetase from Escherichia coli: kinetic mechanism and inhibition by N3-fumaroyl-L-2,3-diaminopropionic derivatives. *Biochemistry* **27**, 2282-2287 (1988).
173 Moraes, G. L., Gomes, G. C., Monteiro de Sousa, P. R., Alves, C. N., Govender, T., Kruger, H. G., Maguire, G. E. M., Lamichhane, G. & Lameira, J. Structural and functional features of enzymes of Mycobacterium tuberculosis peptidoglycan biosynthesis as targets for drug development. *Tuberculosis* **95**, 95-111 (2015).
174 Mengin-Lecreulx, D. & van Heijenoort, J. Characterization of the Essential Gene glmM Encoding Phosphoglucosamine Mutase in Escherichia coli. *Journal of Biological Chemistry* **271**, 32-39 (1996).
175 Mehra-Chaudhary, R., Mick, J. & Beamer, L. J. Crystal Structure of Bacillus anthracis Phosphoglucosamine Mutase, an Enzyme in the Peptidoglycan Biosynthetic Pathway. *Journal of Bacteriology* **193**, 4081-4087 (2011).
176 Raymond, J. B., Mahapatra, S., Crick, D. C. & Pavelka, M. S. Identification of the namH Gene, Encoding the Hydroxylase Responsible for the N-Glycolylation of the Mycobacterial Peptidoglycan. *Journal of Biological Chemistry* **280**, 326-333 (2005).
177 Abrahams, K. A. & Besra, G. S. Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology*, 1-18 (2016).
178 Typas, A., Banzhaf, M., Gross, C. A. & Vollmer, W. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology* **10**, 123 (2011).
179 Chellat, M., Raguž, L. & Riedl, R. *Targeting Antibiotic Resistance*. Vol. 55 (2016).
180 Park, J. T. Uridine-5'-pyrophosphate derivatives: isolation from staphylococcus aureus. *Journal of Biological Chemistry* **194**, 877-884 (1952).
181 Kurosu, M., Mahapatra, S., Narayanasamy, P. & Crick, D. C. Chemoenzymatic synthesis of Park's nucleotide: toward the development of high-throughput screening for MraY inhibitors. *Tetrahedron Letters* **48**, 799-803 (2007).

182 El Zoeiby, A., Sanschagrin, F. & Levesque, R. C. Structure and function of the Mur enzymes: development of novel inhibitors. *Molecular Microbiology* **47**, 1-12 (2003).
183 Smith, C. A. Structure, Function and Dynamics in the mur Family of Bacterial Cell Wall Ligases. *Journal of Molecular Biology* **362**, 640-655 (2006).
184 Squeglia, F., Ruggiero, A. & Berisio, R. Chemistry of Peptidoglycan in Mycobacterium tuberculosis Life Cycle: An off-the-wall Balance of Synthesis and Degradation. *Chemistry – A European Journal* **24**, 2533-2546 (2018).
185 Eniyan, K., Kumar, A., Rayasam, G. V., Perdih, A. & Bajpai, U. Development of a one-pot assay for screening and identification of Mur pathway inhibitors in Mycobacterium tuberculosis. *Scientific Reports* **6**, 35134 (2016).
186 Eniyan, K., Dharavath, S., Vijayan, R., Bajpai, U. & Gourinath, S. Crystal structure of UDP-N-acetylglucosamine-enolpyruvate reductase (MurB) from Mycobacterium tuberculosis. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1866**, 397-406 (2018).
187 Mahapatra, S., Crick, D. C. & Brennan, P. J. Comparison of the UDP-N-Acetylmuramate:l-Alanine Ligase Enzymes from Mycobacterium tuberculosis and Mycobacterium leprae. *Journal of Bacteriology* **182**, 6827-6830 (2000).
188 Barreteau, H., Sosič, I., Turk, S., Humljan, J., Tomašić, T., Zidar, N., Hervé, M., Boniface, A., Peterlin-Mašič, L., Kikelj, D., Mengin-Lecreulx, D., Gobec, S. & Blanot, D. MurD enzymes from different bacteria: Evaluation of inhibitors. *Biochemical Pharmacology* **84**, 625-632 (2012).
189 Basavannacharya, C., Robertson, G., Munshi, T., Keep, N. H. & Bhakta, S. ATP-dependent MurE ligase in Mycobacterium tuberculosis: Biochemical and structural characterisation. *Tuberculosis* **90**, 16-24 (2010).
190 Munshi, T., Gupta, A., Evangelopoulos, D., Guzman, J. D., Gibbons, S., Keep, N. H. & Bhakta, S. Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in Mycobacterium tuberculosis. *PloS one* **8**, e60143-e60143 (2013).
191 Prosser, G. A. & de Carvalho, L. P. S. Kinetic mechanism and inhibition of Mycobacterium tuberculosis d-alanine:d-alanine ligase by the antibiotic d-cycloserine. *The FEBS Journal* **280**, 1150-1166 (2013).
192 Walsh, C. T. Enzymes in the D-alanine branch of bacterial cell wall peptidoglycan assembly. *Journal of Biological Chemistry* **264**, 2393-2396 (1989).
193 Bruning, J. B., Murillo, A. C., Chacon, O., Barletta, R. G. & Sacchettini, J. C. Structure of the Mycobacterium tuberculosis d-Alanine:d-Alanine Ligase, a Target of the Antituberculosis Drug d-Cycloserine. *Antimicrobial Agents and Chemotherapy* **55**, 291-301 (2011).
194 Bouhss, A., Trunkfield, A. E., Bugg, T. D. H. & Mengin-Lecreulx, D. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiology Reviews* **32**, 208-233 (2008).
195 Trunkfield, A. E., Gurcha, S. S., Besra, G. S. & Bugg, T. D. H. Inhibition of Escherichia coli glycosyltransferase MurG and Mycobacterium tuberculosis Gal transferase by uridine-linked transition state mimics. *Bioorganic & Medicinal Chemistry* **18**, 2651-2663 (2010).
196 Vollmer, W. Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiology Reviews* **32**, 287-306 (2008).
197 Brennan, P. J. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis* **83**, 91-97 (2003).
198 Jin, Y., Xin, Y., Zhang, W. & Ma, Y. Mycobacterium tuberculosis Rv1302 and Mycobacterium smegmatis MSMEG___4947 have WecA function and

MSMEG__4947 is required for the growth of M. smegmatis. *FEMS Microbiology Letters* **310**, 54-61 (2010).

199 Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide Endotoxins. *Annual Review of Biochemistry* **71**, 635-700 (2002).

200 Dobrogosz, W. J. Effect of Amino Sugars on Catabolite Repression in *Escherichia coli*. *Journal of Bacteriology* **95**, 578-584 (1968).

201 White, R. J. Control of amino sugar metabolism in Escherichia coli and isolation of mutants unable to degrade amino sugars. *Biochem J* **106**, 847-858 (1968).

202 Freese, E. B., Cole, R. M., Klofat, W. & Freese, E. Growth, Sporulation, and Enzyme Defects of Glucosamine Mutants of *Bacillus subtilis*. *Journal of Bacteriology* **101**, 1046-1062 (1970).

203 Dutka-Malen, S., Mazodier, P. & Badet, B. Molecular cloning and overexpression of the glucosamine synthetase gene from Escherichia coli. *Biochimie* **70**, 287-290 (1988).

204 Walker, J. E., Gay, N. J., Saraste, M. & Eberle, A. N. DNA sequence around the *Escherichia coli* unc operon. Completion of the sequence of a 17 kilobase segment containing *asnA, oriC, unc, glmS* and *phoS*. *Biochemical Journal* **224**, 799-815 (1984).

205 Nilsson, D., Hove-Jensen, B. & Arnvig, K. Primary structure of the tms and prs genes of Bacillus subtilis. *Molecular and General Genetics MGG* **218**, 565-571 (1989).

206 Hove-Jensen, B. Identification of tms-26 as an allele of the gcaD gene, which encodes N-acetylglucosamine 1-phosphate uridyltransferase in Bacillus subtilis. *Journal of bacteriology* **174**, 6852-6856 (1992).

207 Mengin-Lecreulx, D. & van Heijenoort, J. Identification of the glmU gene encoding N-acetylglucosamine-1-phosphate uridyltransferase in Escherichia coli. *Journal of Bacteriology* **175**, 6150-6157 (1993).

208 Mengin-Lecreulx, D. & van Heijenoort, J. Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase activities of Escherichia coli: characterization of the glmU gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-N-acetylglucosamine synthesis. *Journal of Bacteriology* **176**, 5788-5795 (1994).

209 Gehring, A. M., Lees, W. J., Mindiola, D. J., Walsh, C. T. & Brown, E. D. Acetyltransfer Precedes Uridylyltransfer in the Formation of UDP-N-acetylglucosamine in Separable Active Sites of the Bifunctional GlmU Protein of Escherichia coli. *Biochemistry* **35**, 579-585 (1996).

210 Pompeo, F., Bourne, Y., van Heijenoort, J., Fassy, F. & Mengin-Lecreulx, D. Dissection of the Bifunctional Escherichia coli N-Acetylglucosamine-1-phosphate Uridyltransferase Enzyme into Autonomously Functional Domains and Evidence That Trimerization Is Absolutely Required for Glucosamine-1-phosphate Acetyltransferase Activity and Cell Growth. *Journal of Biological Chemistry* **276**, 3833-3839 (2001).

211 Brown, K., Pompeo, F., Dixon, S., Mengin‐Lecreulx, D., Cambillau, C. & Bourne, Y. Crystal structure of the bifunctional *N*‐acetylglucosamine 1‐phosphate uridyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily. *The EMBO Journal* **18**, 4096-4107 (1999).

212 Olsen, L. R. & Roderick, S. L. Structure of the Escherichia coli GlmU Pyrophosphorylase and Acetyltransferase Active Sites. *Biochemistry* **40**, 1913-1921 (2001).

213 Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F. & Bourne, Y. Crystal Structure of Streptococcus pneumoniae N-Acetylglucosamine-1-phosphate

Uridyltransferase Bound to Acetyl-coenzyme A Reveals a Novel Active Site Architecture. *Journal of Biological Chemistry* **276**, 11844-11851 (2001).

214 Mochalkin, I., Lightle, S., Narasimhan, L., Bornemeier, D., Melnick, M., VanderRoest, S. & McDowell, L. Structure of a small-molecule inhibitor complexed with GlmU from Haemophilus influenzae reveals an allosteric binding site. *Protein Science* **17**, 577-582 (2008).

215 Zhang, Z., Bulloch, E. M. M., Bunker, R. D., Baker, E. N. & Squire, C. J. Structure and function of GlmU from Mycobacterium tuberculosis. *Acta Crystallographica Section D* **65**, 275-283 (2009).

216 Bartling, C. M. & Raetz, C. R. H. Steady-State Kinetics and Mechanism of LpxD, the N-Acyltransferase of Lipid A Biosynthesis. *Biochemistry* **47**, 5290-5302 (2008).

217 Beaman, T. W., Blanchard, J. S. & Roderick, S. L. The Conformational Change and Active Site Structure of Tetrahydrodipicolinate N-Succinyltransferase. *Biochemistry* **37**, 10363-10369 (1998).

218 Ballicora, M. A., Iglesias, A. A. & Preiss, J. ADP-Glucose Pyrophosphorylase, a Regulatory Enzyme for Bacterial Glycogen Synthesis. *Microbiology and Molecular Biology Reviews* **67**, 213-225 (2003).

219 Jin, X., Ballicora, M. A., Preiss, J. & Geiger, J. H. Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *The EMBO journal* **24**, 694-704 (2005).

220 Olsen, L. R., Vetting, M. W. & Roderick, S. L. Structure of the E. coli bifunctional GlmU acetyltransferase active site with substrates and products. *Protein Science* **16** (2007).

221 Buurman, E. T., Andrews, B., Gao, N., Hu, J., Keating, T. A., Lahiri, S., Otterbein, L. R., Patten, A. D., Stokes, S. S. & Shapiro, A. B. In Vitro Validation of Acetyltransferase Activity of GlmU as an Antibacterial Target in Haemophilus influenzae. *Journal of Biological Chemistry* **286**, 40734-40742 (2011).

222 Wang, M., Huang, M., Gu, H., Li, S., Ma, Y. & Wang, J. Mutational analysis to identify the residues essential for the acetyltransferase activity of GlmU in Bacillus subtilis. *RSC Advances* **7**, 13858-13867 (2017).

223 Zhou, Y., Xin, Y., Sha, S. & Ma, Y. Kinetic properties of Mycobacterium tuberculosis bifunctional GlmU. *Archives of Microbiology* **193**, 751 (2011).

224 Patin, D., Bayliss, M., Mengin-Lecreulx, D., Oyston, P. & Blanot, D. *Purification and biochemical characterisation of GlmU from Yersinia pestis*. Vol. 197 (2014).

225 Mochalkin, I., Lightle, S., Zhu, Y., Ohren, J. F., Spessard, C., Chirgadze, N. Y., Banotai, C., Melnick, M. & McDowell, L. Characterization of substrate binding and catalysis in the potential antibacterial target N-acetylglucosamine-1-phosphate uridyltransferase (GlmU). *Protein Science* **16**, 2657-2666 (2007).

226 Plumbridge, J. A., Cochet, O., Souza, J. M., Altamirano, M. M., Calcagno, M. L. & Badet, B. Coordinated regulation of amino sugar-synthesizing and -degrading enzymes in Escherichia coli K-12. *Journal of bacteriology* **175**, 4951-4956 (1993).

227 Marshall, S., Bacote, V. & Traxinger, R. R. Discovery of a metabolic pathway mediating glucose-induced desensitization of the glucose transport system. Role of hexosamine biosynthesis in the induction of insulin resistance. *Journal of Biological Chemistry* **266**, 4706-4712 (1991).

228 Broschat, K. O., Gorka, C., Page, J. D., Martin-Berger, C. L., Davies, M. S., Huang, H.-c., Gulve, E. A., Salsgiver, W. J. & Kasten, T. P. Kinetic Characterization of Human Glutamine-fructose-6-phosphate Amidotransferase I: potent feedback inhibition by glucosamine 6-phosphate. *Journal of Biological Chemistry* **277**, 14764-14770 (2002).

229 Peneff, C., Ferrari, P., Charrier, V., Taburet, Y., Monnier, C., Zamboni, V., Winter, J., Harnois, M., Fassy, F. & Bourne, Y. Crystal structures of two human

pyrophosphorylase isoforms in complexes with UDPGlc(Gal)NAc: role of the alternatively spliced insert in the enzyme oligomeric assembly and active site architecture. *The EMBO Journal* **20**, 6191-6202 (2001).

230 Blankenfeldt, W., Asuncion, M., Lam, J. S. & Naismith, J. H. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *The EMBO journal* **19**, 6652-6663 (2000).

231 Wang, J., Liu, X., Liang, Y.-H., Li, L.-F. & Su, X.-D. Acceptor substrate binding revealed by crystal structure of human glucosamine-6-phosphate N-acetyltransferase 1. *FEBS Letters* **582**, 2973-2978 (2008).

232 Pang, H., Koda, Y., Soejima, M. & Kimura, H. Identification of human phosphoglucomutase 3 (PGM3) as N-acetylglucosamine-phosphate mutase (AGM1). PGM3 is equivalent to AGM1. *Annals of Human Genetics* **66**, 139-144 (2002).

233 Muenks, A. G., Stiers, K. M. & Beamer, L. J. Sequence-structure relationships, expression profiles, and disease-associated mutations in the paralogs of phosphoglucomutase 1. *PLOS ONE* **12**, e0183563 (2017).

234 Milewski, S., Gabriel, I. & Olchowy, J. Enzymes of UDP-GlcNAc biosynthesis in yeast. *Yeast* **23**, 1-14 (2006).

235 Watzele, G. & Tanner, W. Cloning of the glutamine:fructose-6-phosphate amidotransferase gene from yeast. Pheromonal regulation of its transcription. *Journal of Biological Chemistry* **264**, 8753-8758 (1989).

236 Raczynska, J., Olchowy, J., Konariev, P. V., Svergun, D. I., Milewski, S. & Rypniewski, W. The Crystal and Solution Studies of Glucosamine-6-phosphate Synthase from Candida albicans. *Journal of Molecular Biology* **372**, 672-688 (2007).

237 Mio, T., Yamada-Okabe, T., Arisawa, M. & Yamada-Okabe, H. Saccharomyces cerevisiae GNA1, an Essential Gene Encoding a Novel Acetyltransferase Involved in UDP-N-acetylglucosamine Synthesis. *Journal of Biological Chemistry* **274**, 424-429 (1999).

238 Peneff, C., Mengin-Lecreulx, D. & Bourne, Y. The Crystal Structures of Apo and Complexed Saccharomyces cerevisiae GNA1 Shed Light on the Catalytic Mechanism of an Amino-sugar N-Acetyltransferase. *Journal of Biological Chemistry* **276**, 16328-16334 (2001).

239 Nishitani, Y., Maruyama, D., Nonaka, T., Kita, A., Fukami, T. A., Mio, T., Yamada-Okabe, H., Yamada-Okabe, T. & Miki, K. Crystal Structures of N-Acetylglucosamine-phosphate Mutase, a Member of the α-d-Phosphohexomutase Superfamily, and Its Substrate and Product Complexes. *Journal of Biological Chemistry* **281**, 19740-19747 (2006).

240 Mio, T., Yabe, T., Arisawa, M. & Yamada-Okabe, H. The Eukaryotic UDP-N-Acetylglucosamine Pyrophosphorylases: gene cloning, protein expression and catalytic mechanism. *Journal of Biological Chemistry* **273**, 14392-14397 (1998).

241 Maruyama, D., Nishitani, Y., Nonaka, T., Kita, A., Fukami, T. A., Mio, T., Yamada-Okabe, H., Yamada-Okabe, T. & Miki, K. Crystal Structure of Uridine-diphospho-N-acetylglucosamine Pyrophosphorylase from Candida albicans and Catalytic Reaction Mechanism. *Journal of Biological Chemistry* **282**, 17221-17230 (2007).

242 Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry Iii, C. E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S. & Barrell, B. G.

Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537 (1998).

243 Camus, J.-C., Pryor, M. J., Médigue, C. & Cole, S. T. Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv. *Microbiology* **148**, 2967-2973 (2002).

244 Zhang, W., Jones, V. C., Scherman, M. S., Mahapatra, S., Crick, D., Bhamidi, S., Xin, Y., McNeil, M. R. & Ma, Y. Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase. *The International Journal of Biochemistry & Cell Biology* **40**, 2560-2571 (2008).

245 Jagtap, P. K. A., Soni, V., Vithani, N., Jhingan, G. D., Bais, V. S., Nandicoori, V. K. & Prakash, B. Substrate-bound Crystal Structures Reveal Features Unique to Mycobacterium tuberculosis N-Acetyl-glucosamine 1-Phosphate Uridyltransferase and a Catalytic Mechanism for Acetyl Transfer. *Journal of Biological Chemistry* **287**, 39524-39537 (2012).

246 Jagtap, P. K. A., Verma, S. K., Vithani, N., Bais, V. S. & Prakash, B. Crystal Structures Identify an Atypical Two-Metal-Ion Mechanism for Uridyltransfer in GlmU: Its Significance to Sugar Nucleotidyl Transferases. *Journal of Molecular Biology* **425**, 1745-1759 (2013).

247 Vithani, N., Ankush Jagtap, P. K., Verma, S. K., Tripathi, R., Awasthi, S., Nair, N. N. & Prakash, B. Mechanism of Mg2+-Accompanied Product Release in Sugar Nucleotidyltransferases. *Structure* **26**, 459-466.e453 (2018).

248 Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Comprehensive identification of conditionally essential genes in mycobacteria. *Proceedings of the National Academy of Sciences* **98**, 12712-12717 (2001).

249 Ravindran, S. Barbara McClintock and the discovery of jumping genes. *Proceedings of the National Academy of Sciences* **109**, 20198-20199 (2012).

250 Hamer, L., DeZwaan, T. M., Montenegro-Chamorro, M. V., Frank, S. A. & Hamer, J. E. Recent advances in large-scale transposon mutagenesis. *Current Opinion in Chemical Biology* **5**, 67-73 (2001).

251 Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology* **48**, 77-84 (2003).

252 Singh, P. & Cole, S. T. Mycobacterium leprae: genes, pseudogenes and genetic diversity. *Future microbiology* **6**, 57-71 (2011).

253 Shiloh, M. U. & DiGiuseppe Champion, P. A. To catch a killer. What can mycobacterial models teach us about Mycobacterium tuberculosis pathogenesis? *Current Opinion in Microbiology* **13**, 86-92 (2010).

254 Soni, V., Upadhayay, S., Suryadevara, P., Samla, G., Singh, A., Yogeeswari, P., Sriram, D. & Nandicoori, V. K. Depletion of M. tuberculosis GlmU from Infected Murine Lungs Effects the Clearance of the Pathogen. *PLOS Pathogens* **11**, e1005235 (2015).

255 DeJesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., Rubin, E. J., Schnappinger, D., Ehrt, S., Fortune, S. M., Sassetti, C. M. & Ioerger, T. R. Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. *mBio* **8**, e02133-02116 (2017).

256 Griffin, J. E., Gawronski, J. D., DeJesus, M. A., Ioerger, T. R., Akerley, B. J. & Sassetti, C. M. High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. *PLOS Pathogens* **7**, e1002251 (2011).

257 Zhang, Y. J., Ioerger, T. R., Huttenhower, C., Long, J. E., Sassetti, C. M., Sacchettini, J. C. & Rubin, E. J. Global Assessment of Genomic Regions

Required for Growth in Mycobacterium tuberculosis. *PLOS Pathogens* **8**, e1002946 (2012).
258 Kinch, M. S., Haynesworth, A., Kinch, S. L. & Hoyer, D. An overview of FDA-approved new molecular entities: 1827–2013. *Drug Discovery Today* **19**, 1033-1039 (2014).
259 Mullard, A. 2018 FDA drug approvals. *Nature Reviews Drug Discovery* **18**, 85-89 (2019).
260 Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R. & Schacht, A. L. How to improve R and D productivity: the pharmaceutical industries grand challenge. *Nature Reviews Drug Discovery* **9**, 203 (2010).
261 Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R and D efficiency. *Nature Reviews Drug Discovery* **11**, 191 (2012).
262 Toure, M. & Crews, C. M. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angewandte Chemie International Edition* **55**, 1966-1973 (2016).
263 Silverstein, A. M. Paul Ehrlich's Passion: The Origins of His Receptor Immunology. *Cellular Immunology* **194**, 213-221 (1999).
264 Rang, H. P. The receptor concept: pharmacology's big idea. *British journal of pharmacology* **147 Suppl 1**, S9-S16 (2006).
265 Bosch, F. & Rosich, L. The contributions of Paul Ehrlich to pharmacology: a tribute on the occasion of the centenary of his Nobel Prize. *Pharmacology* **82**, 171-179 (2008).
266 Gaynes, R. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerging Infectious Diseases* **23**, 849-853 (2017).
267 Ligon, B. L. Penicillin: its discovery and early development. *Seminars in Pediatric Infectious Diseases* **15**, 52-57 (2004).
268 Chain, E., Florey, H. W., Gardner, A. D., Heatley, N. G., Jennings, M. A., Orr-Ewing, J. & Sanders, A. G. Penicillin as a chemotherapeutic agent. *The Lancet* **236**, 226-228 (1940).
269 Fleming, A. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. *British journal of experimental pathology* **10**, 226-236 (1929).
270 Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336 (2016).
271 Borel, J. F., Feurer, C., Magnée, C. & Stähelin, H. Effects of the new anti-lymphocytic peptide cyclosporin A in animals. *Immunology* **32**, 1017-1025 (1977).
272 Shen, B. A New Golden Age of Natural Products Drug Discovery. *Cell* **163**, 1297-1300 (2015).
273 Harvey, A. L., Edrada-Ebel, R. & Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery* **14**, 111 (2015).
274 Meldrum, N. U. & Roughton, F. J. Carbonic anhydrase. Its preparation and properties. *The Journal of physiology* **80**, 113-142 (1933).
275 Schwartz, W. B. The effect of sulfanilamide on salt and water excretion in congestive heart failure. *N Engl J Med* **240**, 173-177 (1949).
276 Supuran, C. T. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nature Reviews Drug Discovery* **7**, 168 (2008).
277 Kenakin, T. Principles: Receptor theory in pharmacology. *Trends in Pharmacological Sciences* **25**, 186-192 (2004).

278 Maehle, A.-H., Prüll, C.-R. & Halliwell, R. F. The emergence of the drug receptor theory. *Nature Reviews Drug Discovery* **1**, 637 (2002).
279 Ahlquist, R. P. A study of the Adrenotropic receptors *American Journal of Physiology-Legacy Content* **153**, 586-600 (1948).
280 Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nature Reviews Drug Discovery* **16**, 829 (2017).
281 Belfield, G. P. & Delaney, S. J. The impact of molecular biology on drug discovery. *Biochemical Society Transactions* **34**, 313-316 (2006).
282 Liu, R., Li, X. & Lam, K. S. Combinatorial chemistry in drug discovery. *Current opinion in chemical biology* **38**, 117-126 (2017).
283 Beeley, L. J. & Duckworth, M. D. The impact of genomics on drug design. *Drug Discovery Today* **1**, 474-480 (1996).
284 Lockhart, D. J. & Winzeler, E. A. Genomics, gene expression and DNA arrays. *Nature* **405**, 827 (2000).
285 Swinney, D. C. The Contribution of Mechanistic Understanding to Phenotypic Screening for First-in-Class Medicines. *Journal of Biomolecular Screening* **18**, 1186-1192 (2013).
286 Macarrón, R. & Hertzberg, R. P. Design and Implementation of High Throughput Screening Assays. *Molecular Biotechnology* **47**, 270-285 (2011).
287 Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V. S., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U. & Sittampalam, G. S. Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery* **10**, 188 (2011).
288 Rees, S., Gribbon, P., Birmingham, K., Janzen, W. P. & Pairaudeau, G. Towards a hit for every target. *Nature Reviews Drug Discovery* **15**, 1 (2015).
289 Baell, J. B. & Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of Medicinal Chemistry* **53**, 2719-2740 (2010).
290 Aldrich, C., Bertozzi, C., Georg, G. I., Kiessling, L., Lindsley, C., Liotta, D., Merz, K. M., Schepartz, A. & Wang, S. The Ecstasy and Agony of Assay Interference Compounds. *ACS Central Science* **3**, 143-147 (2017).
291 Chakravorty, S. J., Chan, J., Greenwood, M. N., Popa-Burke, I., Remlinger, K. S., Pickett, S. D., Green, D. V. S., Fillmore, M. C., Dean, T. W., Luengo, J. I. & Macarrón, R. Nuisance Compounds, PAINS Filters, and Dark Chemical Matter in the GSK HTS Collection. *SLAS DISCOVERY: Advancing Life Sciences R&D* **23**, 532-545 (2018).
292 Hopkins, A. L. & Groom, C. R. The druggable genome. *Nature Reviews Drug Discovery* **1**, 727 (2002).
293 Imming, P., Sinning, C. & Meyer, A. Drugs, their targets and the nature and number of drug targets. *Nature Reviews Drug Discovery* **5**, 821 (2006).
294 Meek, T. D. Inhibitors of HIV-1 Protease. *Journal of Enzyme Inhibition* **6**, 65-98 (1992).
295 Robertson, J. G. Mechanistic Basis of Enzyme-Targeted Drugs. *Biochemistry* **44**, 5561-5571 (2005).
296 Holdgate, G. A., Meek, T. D. & Grimley, R. L. Mechanistic enzymology in drug discovery: a fresh perspective. *Nature Reviews Drug Discovery* **17**, 115 (2017).
297 Copeland, R. A., Pompliano, D. L. & Meek, T. D. Drug–target residence time and its implications for lead optimization. *Nature Reviews Drug Discovery* **5**, 730 (2006).
298 Copeland, R. A. The drug–target residence time model: a 10-year retrospective. *Nature Reviews Drug Discovery* **15**, 87 (2015).

299 Tonge, P. J. Drug–Target Kinetics in Drug Discovery. *ACS Chemical Neuroscience* **9**, 29-39 (2018).
300 Renaud, J.-P., Chung, C.-w., Danielson, U. H., Egner, U., Hennig, M., Hubbard, R. E. & Nar, H. Biophysics in drug discovery: impact, challenges and opportunities. *Nature Reviews Drug Discovery* **15**, 679 (2016).
301 Schramm, V. L. Transition States, Analogues, and Drug Development. *ACS Chemical Biology* **8**, 71-81 (2013).
302 Dreyer, G. B., Metcalf, B. W., Tomaszek, T. A., Carr, T. J., Chandler, A. C., Hyland, L., Fakhoury, S. A., Magaard, V. W., Moore, M. L. & Strickler, J. E. Inhibition of human immunodeficiency virus 1 protease in vitro: rational design of substrate analogue inhibitors. *Proc Natl Acad Sci U S A* **86**, 9752-9756 (1989).
303 Staessen, J. A., Li, Y. & Richart, T. Oral renin inhibitors. *The Lancet* **368**, 1449-1456 (2006).
304 Blundell, T. L., Jhoti, H. & Abell, C. High-throughput crystallography for lead discovery in drug design. *Nature Reviews Drug Discovery* **1**, 45-54 (2002).
305 Jhoti, H., Cleasby, A., Verdonk, M. & Williams, G. Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Current Opinion in Chemical Biology* **11**, 485-493 (2007).
306 Silvestre, H. L., Blundell, T. L., Abell, C. & Ciulli, A. Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery. *Proceedings of the National Academy of Sciences* **110**, 12984-12989 (2013).
307 Brenner, S. & Lerner, R. A. Encoded combinatorial chemistry. *Proceedings of the National Academy of Sciences* **89**, 5381-5383 (1992).
308 Mullard, A. DNA tags help the hunt for drugs. *Nature* **530**, 367 (2016).
309 Clark, M. A., Acharya, R. A., Arico-Muendel, C. C., Belyanskaya, S. L., Benjamin, D. R., Carlson, N. R., Centrella, P. A., Chiu, C. H., Creaser, S. P., Cuozzo, J. W., Davie, C. P., Ding, Y., Franklin, G. J., Franzen, K. D., Gefter, M. L., Hale, S. P., Hansen, N. J. V., Israel, D. I., Jiang, J., Kavarana, M. J., Kelley, M. S., Kollmann, C. S., Li, F., Lind, K., Mataruse, S., Medeiros, P. F., Messer, J. A., Myers, P., O'Keefe, H., Oliff, M. C., Rise, C. E., Satz, A. L., Skinner, S. R., Svendsen, J. L., Tang, L., van Vloten, K., Wagner, R. W., Yao, G., Zhao, B. & Morgan, B. A. Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nature Chemical Biology* **5**, 647 (2009).
310 Machutta, C. A., Kollmann, C. S., Lind, K. E., Bai, X., Chan, P. F., Huang, J., Ballell, L., Belyanskaya, S., Besra, G. S., Barros-Aguirre, D., Bates, R. H., Centrella, P. A., Chang, S. S., Chai, J., Choudhry, A. E., Coffin, A., Davie, C. P., Deng, H., Deng, J., Ding, Y., Dodson, J. W., Fosbenner, D. T., Gao, E. N., Graham, T. L., Graybill, T. L., Ingraham, K., Johnson, W. P., King, B. W., Kwiatkowski, C. R., Lelièvre, J., Li, Y., Liu, X., Lu, Q., Lehr, R., Mendoza-Losana, A., Martin, J., McCloskey, L., McCormick, P., O'Keefe, H. P., O'Keeffe, T., Pao, C., Phelps, C. B., Qi, H., Rafferty, K., Scavello, G. S., Steiginga, M. S., Sundersingh, F. S., Sweitzer, S. M., Szewczuk, L. M., Taylor, A., Toh, M. F., Wang, J., Wang, M., Wilkins, D. J., Xia, B., Yao, G., Zhang, J., Zhou, J., Donahue, C. P., Messer, J. A., Holmes, D., Arico-Muendel, C. C., Pope, A. J., Gross, J. W. & Evindar, G. Prioritizing multiple therapeutic targets in parallel using automated DNA-encoded library screening. *Nature communications* **8**, 16081-16081 (2017).
311 Valler, M. J. & Green, D. Diversity screening versus focussed screening in drug discovery. *Drug Discovery Today* **5**, 286-293 (2000).
312 Bradley, A. R., Echalier, A., Fairhead, M., Strain-Damerell, C., Brennan, P., Bullock, A. N., Burgess-Brown, N. A., Carpenter, E. P., Gileadi, O., Marsden, B.

D., Lee, W. H., Yue, W., Bountra, C. & von Delft, F. The SGC beyond structural genomics: redefining the role of 3D structures by coupling genomic stratification with fragment-based discovery. *Essays in biochemistry* **61**, 495-503 (2017).

313 Drewry, D. H., Willson, T. M. & Zuercher, W. J. Seeding Collaborations to Advance Kinase Science with the GSK Published Kinase Inhibitor Set (PKIS). *Current Topics in Medicinal Chemistry* **14**, 340-342 (2014).

314 Kutchukian, P. S., Warren, L., Magliaro, B. C., Amoss, A., Cassaday, J. A., O'Donnell, G., Squadroni, B., Zuck, P., Pascarella, D., Culberson, J. C., Cooke, A. J., Hurzy, D., Schlegel, K.-A. S., Thomson, F., Johnson, E. N., Uebele, V. N., Hermes, J. D., Parmentier-Batteur, S. & Finley, M. Iterative Focused Screening with Biological Fingerprints Identifies Selective Asc-1 Inhibitors Distinct from Traditional High Throughput Screening. *ACS Chemical Biology* **12**, 519-527 (2017).

315 Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nature Reviews Drug Discovery* **10**, 507 (2011).

316 Zheng, W., Thorne, N. & McKew, J. C. Phenotypic screens as a renewed approach for drug discovery. *Drug Discovery Today* **18**, 1067-1073 (2013).

317 Couper, K. N., Blount, D. G. & Riley, E. M. IL-10: The Master Regulator of Immunity to Infection. *The Journal of Immunology* **180**, 5771-5777 (2008).

318 Aderem, A. & Underhill, D. M. Mechanisms of phagocytosis in macrophages *Annual Review of Immunology* **17**, 593-623 (1999).

319 Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology* **4**, 682 (2008).

320 Moffat, J. G. Turning the Light On in the Phenotypic Drug Discovery Black Box. *Cell Chemical Biology* **24**, 545-547 (2017).

321 Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nature Reviews Drug Discovery* **16**, 531 (2017).

322 Drewes, G. & Knapp, S. Chemoproteomics and Chemical Probes for Target Discovery. *Trends in Biotechnology* **36**, 1275-1286 (2018).

323 Freiberg, C. & Brötz-Oesterhelt, H. Functional genomics in antibacterial drug discovery. *Drug Discovery Today* **10**, 927-935 (2005).

324 Vincent, I. M. & Barrett, M. P. Metabolomic-Based Strategies for Anti-Parasite Drug Discovery. *Journal of Biomolecular Screening* **20**, 44-55 (2015).

325 Mirguet, O., Lamotte, Y., Donche, F., Toum, J., Gellibert, F., Bouillot, A., Gosmini, R., Nguyen, V.-L., Delannée, D., Seal, J., Blandel, F., Boullay, A.-B., Boursier, E., Martin, S., Brusq, J.-M., Krysa, G., Riou, A., Tellier, R., Costaz, A., Huet, P., Dudit, Y., Trottet, L., Kirilovsky, J. & Nicodeme, E. From ApoA1 upregulation to BET family bromodomain inhibition: Discovery of I-BET151. *Bioorganic & Medicinal Chemistry Letters* **22**, 2963-2967 (2012).

326 Chung, C.-w., Coste, H., White, J. H., Mirguet, O., Wilde, J., Gosmini, R. L., Delves, C., Magny, S. M., Woodward, R., Hughes, S. A., Boursier, E. V., Flynn, H., Bouillot, A. M., Bamborough, P., Brusq, J.-M. G., Gellibert, F. J., Jones, E. J., Riou, A. M., Homes, P., Martin, S. L., Uings, I. J., Toum, J., Clément, C. A., Boullay, A.-B., Grimley, R. L., Blandel, F. M., Prinjha, R. K., Lee, K., Kirilovsky, J. & Nicodeme, E. Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. *Journal of Medicinal Chemistry* **54**, 3827-3838 (2011).

327 Vane, J. R. Nobel lecture. Adventures and excursions in bioassay--the stepping stones to prostacyclin. *Postgraduate medical journal* **59**, 743-758 (1983).

328 Vane, J. R. & Botting, R. M. The mechanism of action of aspirin. *Thrombosis Research* **110**, 255-258 (2003).

329 Clatworthy, A. E., Romano, K. P. & Hung, D. T. Whole-organism phenotypic screening for anti-infectives promoting host health. *Nature Chemical Biology* **14**, 331-341 (2018).

330 Plouffe, D., Brinker, A., McNamara, C., Henson, K., Kato, N., Kuhen, K., Nagle, A., Adrián, F., Matzen, J. T., Anderson, P., Nam, T.-g., Gray, N. S., Chatterjee, A., Janes, J., Yan, S. F., Trager, R., Caldwell, J. S., Schultz, P. G., Zhou, Y. & Winzeler, E. A. *In silico* activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proceedings of the National Academy of Sciences* **105**, 9059-9064 (2008).

331 Gamo, F.-J., Sanz, L. M., Vidal, J., de Cozar, C., Alvarez, E., Lavandera, J.-L., Vanderwall, D. E., Green, D. V. S., Kumar, V., Hasan, S., Brown, J. R., Peishoff, C. E., Cardon, L. R. & Garcia-Bustos, J. F. Thousands of chemical starting points for antimalarial lead identification. *Nature* **465**, 305 (2010).

332 Calderón, F., Barros, D., Bueno, J. M., Coterón, J. M., Fernández, E., Gamo, F. J., Lavandera, J. L., León, M. L., Macdonald, S. J. F., Mallo, A., Manzano, P., Porras, E., Fiandor, J. M. & Castro, J. An Invitation to Open Innovation in Malaria Drug Discovery: 47 Quality Starting Points from the TCAMS. *ACS Medicinal Chemistry Letters* **2**, 741-746 (2011).

333 Guiguemde, W. A., Shelat, A. A., Bouck, D., Duffy, S., Crowther, G. J., Davis, P. H., Smithson, D. C., Connelly, M., Clark, J., Zhu, F., Jiménez-Díaz, M. B., Martinez, M. S., Wilson, E. B., Tripathi, A. K., Gut, J., Sharlow, E. R., Bathurst, I., Mazouni, F. E., Fowble, J. W., Forquer, I., McGinley, P. L., Castro, S., Angulo-Barturen, I., Ferrer, S., Rosenthal, P. J., DeRisi, J. L., Sullivan, D. J., Lazo, J. S., Roos, D. S., Riscoe, M. K., Phillips, M. A., Rathod, P. K., Van Voorhis, W. C., Avery, V. M. & Guy, R. K. Chemical genetics of Plasmodium falciparum. *Nature* **465**, 311 (2010).

334 Ballell, L., Strange, M., Cammack, N., Fairlamb, A. H. & Borysiewicz, L. Open Lab as a source of hits and leads against tuberculosis, malaria and kinetoplastid diseases. *Nature Reviews Drug Discovery* **15**, 292 (2016).

335 Ballell, L., Bates, R. H., Young, R. J., Alvarez-Gomez, D., Alvarez-Ruiz, E., Barroso, V., Blanco, D., Crespo, B., Escribano, J., González, R., Lozano, S., Huss, S., Santos-Villarejo, A., Martín-Plaza, J. J., Mendoza, A., Rebollo-Lopez, M. J., Remuiñan-Blanco, M., Lavandera, J. L., Pérez-Herran, E., Gamo-Benito, F. J., García-Bustos, J. F., Barros, D., Castro, J. P. & Cammack, N. Fueling Open-Source Drug Discovery: 177 Small-Molecule Leads against Tuberculosis. *ChemMedChem* **8**, 313-321 (2013).

336 Rebollo-Lopez, M. J., Lelièvre, J., Alvarez-Gomez, D., Castro-Pichel, J., Martínez-Jiménez, F., Papadatos, G., Kumar, V., Colmenarejo, G., Mugumbate, G., Hurle, M., Barroso, V., Young, R. J., Martinez-Hoyos, M., González del Río, R., Bates, R. H., Lopez-Roman, E. M., Mendoza-Losana, A., Brown, J. R., Alvarez-Ruiz, E., Marti-Renom, M. A., Overington, J. P., Cammack, N., Ballell, L. & Barros-Aguire, D. Release of 50 new, drug-like compounds and their computational target predictions for open source anti-tubercular drug discovery. *PLOS ONE* **10**, e0142293 (2015).

337 Peña, I., Pilar Manzano, M., Cantizani, J., Kessler, A., Alonso-Padilla, J., Bardera, A. I., Alvarez, E., Colmenarejo, G., Cotillo, I., Roquero, I., de Dios-Anton, F., Barroso, V., Rodriguez, A., Gray, D. W., Navarro, M., Kumar, V., Sherstnev, A., Drewry, D. H., Brown, J. R., Fiandor, J. M. & Julio Martin, J. New Compound Sets Identified from High Throughput Phenotypic Screening Against Three Kinetoplastid Parasites: An Open Resource. *Scientific Reports* **5**, 8771 (2015).

338 Andries, K., Verhasselt, P., Guillemont, J., Göhlmann, H. W. H., Neefs, J.-M., Winkler, H., Van Gestel, J., Timmerman, P., Zhu, M., Lee, E., Williams, P., de

Chaffoy, D., Huitric, E., Hoffner, S., Cambau, E., Truffot-Pernot, C., Lounis, N. & Jarlier, V. A Diarylquinoline Drug Active on the ATP Synthase of *Mycobacterium tuberculosis*. *Science* **307**, 223-227 (2005).

339 Abrahams, K. A., Chung, C.-w., Ghidelli-Disse, S., Rullas, J., Rebollo-López, M. J., Gurcha, S. S., Cox, J. A. G., Mendoza, A., Jiménez-Navarro, E., Martínez-Martínez, M. S., Neu, M., Shillings, A., Homes, P., Argyrou, A., Casanueva, R., Loman, N. J., Moynihan, P. J., Lelièvre, J., Selenski, C., Axtman, M., Kremer, L., Bantscheff, M., Angulo-Barturen, I., Izquierdo, M. C., Cammack, N. C., Drewes, G., Ballell, L., Barros, D., Besra, G. S. & Bates, R. H. Identification of KasA as the cellular target of an anti-tubercular scaffold. *Nature Communications* **7**, 12581 (2016).

340 Abrahams, K. A., Cox, J. A. G., Fütterer, K., Rullas, J., Ortega-Muro, F., Loman, N. J., Moynihan, P. J., Pérez-Herrán, E., Jiménez, E., Esquivias, J., Barros, D., Ballell, L., Alemparte, C. & Besra, G. S. Inhibiting mycobacterial tryptophan synthase by targeting the inter-subunit interface. *Scientific Reports* **7**, 9430 (2017).

341 Cox, J. A. G., Mugumbate, G., Del Peral, L. V.-G., Jankute, M., Abrahams, K. A., Jervis, P., Jackenkroll, S., Perez, A., Alemparte, C., Esquivias, J., Lelièvre, J., Ramon, F., Barros, D., Ballell, L. & Besra, G. S. Novel inhibitors of Mycobacterium tuberculosis GuaB2 identified by a target based high-throughput phenotypic screen. *Scientific Reports* **6**, 38986 (2016).

342 Katsuno, K., Burrows, J. N., Duncan, K., van Huijsduijnen, R. H., Kaneko, T., Kita, K., Mowbray, C. E., Schmatz, D., Warner, P. & Slingsby, B. T. Hit and lead criteria in drug discovery for infectious diseases of the developing world. *Nature Reviews Drug Discovery* **14**, 751 (2015).

343 Abraham, E. P. & Chain, E. An Enzyme from Bacteria able to Destroy Penicillin. *Nature* **146**, 837 (1940).

344 Kumarasamy, K. K., Toleman, M. A., Walsh, T. R., Bagaria, J., Butt, F., Balakrishnan, R., Chaudhary, U., Doumith, M., Giske, C. G., Irfan, S., Krishnan, P., Kumar, A. V., Maharjan, S., Mushtaq, S., Noorie, T., Paterson, D. L., Pearson, A., Perry, C., Pike, R., Rao, B., Ray, U., Sarma, J. B., Sharma, M., Sheridan, E., Thirunarayan, M. A., Turton, J., Upadhyay, S., Warner, M., Welfare, W., Livermore, D. M. & Woodford, N. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *The Lancet Infectious Diseases* **10**, 597-602 (2010).

345 Fischbach, M. A. & Walsh, C. T. Antibiotics for Emerging Pathogens. *Science* **325**, 1089-1093 (2009).

346 Hofer, U. The cost of antimicrobial resistance. *Nature Reviews Microbiology* **17**, 3-3 (2019).

347 Cassini, A., Högberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., Colomb-Cotinat, M., Kretzschmar, M. E., Devleesschauwer, B., Cecchini, M., Ouakrim, D. A., Oliveira, T. C., Struelens, M. J., Suetens, C., Monnet, D. L., Strauss, R., Mertens, K., Struyf, T., Catry, B., Latour, K., Ivanov, I. N., Dobreva, E. G., Tambic Andraševic, A., Soprek, S., Budimir, A., Paphitou, N., Žemlicková, H., Schytte Olsen, S., Wolff Sönksen, U., Märtin, P., Ivanova, M., Lyytikäinen, O., Jalava, J., Coignard, B., Eckmanns, T., Abu Sin, M., Haller, S., Daikos, G. L., Gikas, A., Tsiodras, S., Kontopidou, F., Tóth, Á., Hajdu, Á., Guólaugsson, Ó., Kristinsson, K. G., Murchan, S., Burns, K., Pezzotti, P., Gagliotti, C., Dumpis, U., Liuimiene, A., Perrin, M., Borg, M. A., de Greeff, S. C., Monen, J. C. M., Koek, M. B. G., Elstrøm, P., Zabicka, D., Deptula, A., Hryniewicz, W., Caniça, M., Nogueira, P. J., Fernandes, P. A., Manageiro, V., Popescu, G. A., Serban, R. I., Schréterová, E., Litvová, S., Štefkovicová, M., Kolman, J., Klavs, I., Korošec, A., Aracil, B., Asensio, A., Pérez-Vázquez, M.,

Billström, H., Larsson, S., Reilly, J. S., Johnson, A. & Hopkins, S. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *The Lancet Infectious Diseases* **19**, 56-66 (2019).

348 Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature Reviews Drug Discovery* **6**, 29 (2006).

349 Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I. & Miller, A. A. ESKAPEing the labyrinth of antibacterial discovery. *Nature Reviews Drug Discovery* **14**, 529 (2015).

350 Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J. F. M. L., Sekiguchi, J., Sekowska, A., Séror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U. & Ogasawara, N. Essential Bacillus subtilis genes. *Proceedings of the National Academy of Sciences* **100**, 4678-4683 (2003).

351 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. & Mori, H. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* **2**, 2006.0008-2006.0008 (2006).

352 Forsyth, R. A., Haselbeck, R. J., Ohlsen, K. L., Yamamoto, R. T., Xu, H., Trawick, J. D., Wall, D., Wang, L., Brown-Driver, V., Froelich, J. M., C., K. G., King, P., McCarthy, M., Malone, C., Misiner, B., Robbins, D., Tan, Z., Zhu, Z.-y., Carr, G., Mosca, D. A., Zamudio, C., Foulkes, J. G. & Zyskind, J. W. A genome-wide strategy for the identification of essential genes in Staphylococcus aureus. *Molecular Microbiology* **43**, 1387-1400 (2002).

353 Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M. V. & Manoil, C. Comprehensive transposon mutant library of Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences* **100**, 14339-14344 (2003).

354 Jain, P., Hsu, T., Arai, M., Biermann, K., Thaler, D. S., Nguyen, A., González, P. A., Tufariello, J. M., Kriakov, J., Chen, B., Larsen, M. H. & Jacobs, W. R. Specialized Transduction Designed for Precise High-Throughput Unmarked Deletions in Mycobacterium tuberculosis. *mBio* **5**, e01245-01214 (2014).

355 D'Elia, M. A., Pereira, M. P. & Brown, E. D. Are essential genes really essential? *Trends in Microbiology* **17**, 433-438 (2009).

356 Neuhaus, F. C. & Lynch, J. L. The Enzymatic Synthesis of D-Alanyl-D-alanine. III. On the Inhibition of D-Alanyl-D-alanine Synthetase by the Antibiotic D-Cycloserine*. *Biochemistry* **3**, 471-480 (1964).
357 Lambert, M. P. & Neuhaus, F. C. Mechanism of d-Cycloserine Action: Alanine Racemase from Escherichia coli. *Journal of Bacteriology* **110**, 978-987 (1972).
358 Prosser, G. A. & de Carvalho, L. P. S. Reinterpreting the Mechanism of Inhibition of Mycobacterium tuberculosis d-Alanine:d-Alanine Ligase by d-Cycloserine. *Biochemistry* **52**, 7145-7149 (2013).
359 Batson, S., de Chiara, C., Majce, V., Lloyd, A. J., Gobec, S., Rea, D., Fülöp, V., Thoroughgood, C. W., Simmons, K. J., Dowson, C. G., Fishwick, C. W. G., de Carvalho, L. P. S. & Roper, D. I. Inhibition of D-Ala:D-Ala ligase through a phosphorylated form of the antibiotic D-cycloserine. *Nature Communications* **8**, 1939 (2017).
360 WHO. Global Investments in Tuberculosis Research and Development: Past, Present, and Future. 80 (2016).
361 Partnership, S. T. Tuberculosis Research Funding Trends. *Treatment Action Group* (2018).
362 Murray, J. F. A Century of Tuberculosis. *American Journal of Respiratory and Critical Care Medicine* **169**, 1181-1186 (2004).
363 Luca, S. & Mihaescu, T. History of BCG Vaccine. *Maedica (Buchar)* **8**, 53-58 (2013).
364 Behr, M. A., Warren, S. A., Salamon, H., Hopewell, P. C., de Leon, A. P., Daley, C. L. & Small, P. M. Transmission of Mycobacterium tuberculosis from patients smear-negative for acid-fast bacilli. *The Lancet* **353**, 444-449 (1999).
365 Plotkin, S. History of vaccination. *Proceedings of the National Academy of Sciences* **111**, 12283-12287 (2014).
366 Ebrahim, S. & Ferrie, J. E. One hundred years of Medical Research – the UK Medical Research Council Centenary. *International Journal of Epidemiology* **42**, 930-931 (2013).
367 Crofton, J. The MRC randomized trial of streptomycin and its legacy: a view from the clinical front line. *Journal of the Royal Society of Medicine* **99**, 531-534 (2006).
368 Villemagne, B., Crauste, C., Flipo, M., Baulard, A. R., Déprez, B. & Willand, N. Tuberculosis: The drug development pipeline at a glance. *European Journal of Medicinal Chemistry* **51**, 1-16 (2012).
369 Holme, C. Tuberculosis: story of medical failure. *BMJ (Clinical research ed.)* **317**, 1260-1260 (1998).
370 Kaufmann, S. H. E. & Parida, S. K. Changing funding patterns in tuberculosis. *Nature Medicine* **13**, 299 (2007).
371 Dolin, P. J., Raviglione, M. C. & Kochi, A. Global tuberculosis incidence and mortality during 1990-2000. *Bulletin of the World Health Organization* **72**, 213-220 (1994).
372 Tornheim, J. A. & Dooley, K. E. The Global Landscape of Tuberculosis Therapeutics. *Annual Review of Medicine* **70**, 105-120 (2019).
373 Mahajan, R. Bedaquiline: First FDA-approved tuberculosis drug in 40 years. *Int J Appl Basic Med Res* **3**, 1-2 (2013).
374 Xavier, A. S. & Lakshmanan, M. Delamanid: A new armor in combating drug-resistant tuberculosis. *Journal of pharmacology & pharmacotherapeutics* **5**, 222-224 (2014).
375 Lienhardt, C., Jaramillo, E., Raviglione, M., Spigelman, M., Hafner, R., Hoelscher, M., Zumla, A. & Gheuens, J. New Drugs for the Treatment of Tuberculosis: Needs, Challenges, Promise, and Prospects for the Future. *The Journal of Infectious Diseases* **205**, S241-S249 (2012).

376 Kundu, S., Biukovic, G., Grüber, G. & Dick, T. Bedaquiline Targets the ε Subunit of Mycobacterial F-ATP Synthase. *Antimicrobial Agents and Chemotherapy* **60**, 6977-6979 (2016).
377 Riccardi, G., Pasca, M. R., Chiarelli, L. R., Manina, G., Mattevi, A. & Binda, C. The DprE1 enzyme, one of the most vulnerable targets of Mycobacterium tuberculosis. *Applied Microbiology and Biotechnology* **97**, 8841-8848 (2013).
378 Li, W., Upadhyay, A., Fontes, F. L., North, E. J., Wang, Y., Crans, D. C., Grzegorzewicz, A. E., Jones, V., Franzblau, S. G., Lee, R. E., Crick, D. C. & Jackson, M. Novel insights into the mechanism of inhibition of MmpL3, a target of multiple pharmacophores in Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy* **58**, 6413-6423 (2014).
379 Stanley, S. A., Grant, S. S., Kawate, T., Iwase, N., Shimizu, M., Wivagg, C., Silvis, M., Kazyanskaya, E., Aquadro, J., Golas, A., Fitzgerald, M., Dai, H., Zhang, L. & Hung, D. T. Identification of Novel Inhibitors of M. tuberculosis Growth Using Whole Cell Based High-Throughput Screening. *ACS Chemical Biology* **7**, 1377-1384 (2012).
380 Grzegorzewicz, A. E., Pham, H., Gundi, V. A. K. B., Scherman, M. S., North, E. J., Hess, T., Jones, V., Gruppo, V., Born, S. E. M., Korduláková, J., Chavadi, S. S., Morisseau, C., Lenaerts, A. J., Lee, R. E., McNeil, M. R. & Jackson, M. Inhibition of mycolic acid transport across the Mycobacterium tuberculosis plasma membrane. *Nature Chemical Biology* **8**, 334 (2012).
381 Pereira, M. P., Blanchard, J. E., Murphy, C., Roderick, S. L. & Brown, E. D. High-Throughput Screening Identifies Novel Inhibitors of the Acetyltransferase Activity of Escherichia coli GlmU. *Antimicrobial Agents and Chemotherapy* **53**, 2306-2311 (2009).
382 Rani, C., Mehra, R., Sharma, R., Chib, R., Wazir, P., Nargotra, A. & Khan, I. A. High-throughput screen identifies small molecule inhibitors targeting acetyltransferase activity of Mycobacterium tuberculosis GlmU. *Tuberculosis* **95**, 664-677 (2015).
383 Sharma, R., Rani, C., Mehra, R., Nargotra, A., Chib, R., Rajput, V. S., Kumar, S., Singh, S., Sharma, P. R. & Khan, I. A. Identification and characterization of novel small molecule inhibitors of the acetyltransferase activity of Escherichia coli N-acetylglucosamine-1-phosphate-uridyltransferase/glucosamine-1-phosphate-acetyltransferase (GlmU). *Applied Microbiology and Biotechnology* **100**, 3071-3085 (2016).
384 Green, O. M., McKenzie, A. R., Shapiro, A. B., Otterbein, L., Ni, H., Patten, A., Stokes, S., Albert, R., Kawatkar, S. & Breed, J. Inhibitors of acetyltransferase domain of N-acetylglucosamine-1-phosphate-uridyltransferase/glucosamine-1-phosphate-acetyltransferase (GlmU). Part 1: Hit to lead evaluation of a novel arylsulfonamide series. *Bioorganic & Medicinal Chemistry Letters* **22**, 1510-1519 (2012).
385 Doig, P., Boriack-Sjodin, P. A., Dumas, J., Hu, J., Itoh, K., Johnson, K., Kazmirski, S., Kinoshita, T., Kuroda, S., Sato, T.-o., Sugimoto, K., Tohyama, K., Aoi, H., Wakamatsu, K. & Wang, H. Rational design of inhibitors of the bacterial cell wall synthetic enzyme GlmU using virtual screening and lead-hopping. *Bioorganic & Medicinal Chemistry* **22**, 6256-6269 (2014).
386 Larsen, Nicholas A., Nash, Tory J., Morningstar, M., Shapiro, Adam B., Joubran, C., Blackett, Carolyn J., Patten, Arthur D., Boriack-Sjodin, P. A. & Doig, P. An aminoquinazoline inhibitor of the essential bacterial cell wall synthetic enzyme GlmU has a unique non-protein-kinase-like binding mode. *Biochemical Journal* **446**, 405-413 (2012).
387 Silver, L. L. Appropriate Targets for Antibacterial Drugs. *Cold Spring Harbor Perspectives in Medicine* **6** (2016).

388 Silver, L. L. Challenges of antibacterial discovery. *Clinical microbiology reviews* **24**, 71-109 (2011).
389 Singh, S. B., Young, K. & Silver, L. L. What is an “ideal” antibiotic? Discovery challenges and path forward. *Biochemical Pharmacology* **133**, 63-73 (2017).
390 Morais, L. L., Yuasa, H., Bennis, K., Ripoche, I. & Auzanneau, F.-I. Chemoenzymatic synthesis of thio-nod factor intermediates — Enzymatic transfer of glucosamine on thiochitobiose derivatives. *Canadian Journal of Chemistry* **84**, 587-596 (2006).
391 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254 (1976).
392 Grassetti, D. R. & Murray, J. F. Determination of sulfhydryl groups with 2,2′ - or 4,4′ -dithiodipyridine. *Archives of Biochemistry and Biophysics* **119**, 41-49 (1967).
393 Miller-Martini, D. M., Hua, S. & Horowitz, P. M. Cysteine 254 can cooperate with active site cysteine 247 in reactivation of 5,5'-dithiobis(2-nitrobenzoic acid)-inactivated rhodanese as determined by site-directed mutagenesis. *Journal of Biological Chemistry* **269**, 12414-12418 (1994).
394 Moser, M., Schneider, R., Behnke, T., Schneider, T., Falkenhagen, J. & Resch-Genger, U. Ellman’s and Aldrithiol Assay as Versatile and Complementary Tools for the Quantification of Thiol Groups and Ligands on Nanomaterials. *Analytical Chemistry* **88**, 8624-8631 (2016).
395 Ellman, G. L. Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics* **82**, 70-77 (1959).
396 Vazquez, M. J., Rodriguez, B., Zapatero, C. & Tew, D. G. Determination of phosphate in nanomolar range by an enzyme-coupling fluorescent method. *Analytical Biochemistry* **320**, 292-298 (2003).
397 Brune, M., Hunter, J. L., Corrie, J. E. T. & Webb, M. R. Direct, Real-Time Measurement of Rapid Inorganic Phosphate Release Using a Novel Fluorescent Probe and Its Application to Actomyosin Subfragment 1 ATPase. *Biochemistry* **33**, 8262-8271 (1994).
398 Ellis, K. J. & Morrison, J. F. in *Methods in Enzymology* Vol. 87 (ed Daniel L. Purich) 405-426 (Academic Press, 1982).
399 Zhang, J.-H., Chung, T. D. Y. & Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4**, 67-73 (1999).
400 Beckonert, O., Keun, H. C., Ebbels, T. M. D., Bundy, J., Holmes, E., Lindon, J. C. & Nicholson, J. K. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols* **2**, 2692 (2007).
401 Kay, L., Keifer, P. & Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society* **114**, 10663-10665 (1992).
402 Baryshnikova, O. K., Williams, T. C. & Sykes, B. D. Internal pH indicators for biomolecular NMR. *Journal of Biomolecular NMR* **41**, 5-7 (2008).
403 Oregioni, A., Stieglitz, B., Kelly, G., Rittinger, K. & Frenkiel, T. Determination of the pKa of the N-terminal amino group of ubiquitin by NMR. *Scientific Reports* **7**, 43748 (2017).
404 Derome, A. E. & Williamson, M. P. Rapid-pulsing artifacts in double-quantum-filtered COSY. *Journal of Magnetic Resonance (1969)* **88**, 177-185 (1990).
405 Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas', A. F. & Gilmanshin, R. I. Study of the “molten globule” intermediate state in

protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**, 119-128 (1991).
406 Senisterra, G. A., Markin, E., Yamazaki, K., Hui, R., Vedadi, M. & Awrey, D. E. Screening for Ligands Using a Generic and High-Throughput Light-Scattering-Based Assay. *Journal of Biomolecular Screening* **11**, 940-948 (2006).
407 Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protocols* **2**, 2212-2221 (2007).
408 Lo, M.-C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M. & Ellestad, G. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical Biochemistry* **332**, 153-159 (2004).
409 Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols* **1**, 2876-2890 (2006).
410 McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta crystallographica. Section F, Structural biology communications* **70**, 2-20 (2013).
411 Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr* **69**, 1260-1273 (2013).
412 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).
413 Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**, 213-221 (2010).
414 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **66**, 486-501 (2010).
415 Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallographica Section D* **55**, 191-205 (1999).
416 Schrodinger, L. The PyMOL Molecular Graphics System, Version 1.3r1. (2010).
417 Skoog, D. A. & West, D. M. *Fundamentals of Analytical Chemistry*. (Saunders College Publishing, 1982).
418 Sieber, V., Jurnak, F. & Moe, G. R. Circular dichroism of the parallel β helical proteins pectate lyase C and E. *Proteins: Structure, Function, and Bioinformatics* **23**, 32-37 (1995).
419 Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1751**, 119-139 (2005).
420 Winther, J. R. & Thorpe, C. Quantification of thiols and disulfides. *Biochimica et biophysica acta* **1840**, 838-846 (2014).
421 Riener, C. K., Kada, G. & Gruber, H. J. Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4′ -dithiodipyridine. *Analytical and Bioanalytical Chemistry* **373**, 266-276 (2002).
422 Wright, G. D. Aminoglycoside-modifying enzymes. *Current Opinion in Microbiology* **2**, 499-503 (1999).
423 Strelow, J. M. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS DISCOVERY: Advancing Life Sciences R&D* **22**, 3-20 (2017).

424 Olsen, L. R., Vetting, M. W. & Roderick, S. L. Structure of the E. coli bifunctional GlmU acetyltransferase active site with substrates and products. *Protein Science : A Publication of the Protein Society* **16**, 1230-1235 (2007).
425 Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-5467 (1977).
426 Baell, J. & Walters, M. *Chemistry: Chemical con artists foil drug discovery*. Vol. 513 (2014).
427 Ballell, L., Bates, R. H., Young, R. J., Alvarez-Gomez, D., Alvarez-Ruiz, E., Barroso, V., Blanco, D., Crespo, B., Escribano, J., Gonzalez, R., Lozano, S., Huss, S., Santos-Villarejo, A., Martin-Plaza, J. J., Mendoza, A., Rebollo-Lopez, M. J., emuian-Blanco, M., Perez-Heran, E., Javier, F., rcia-Bustos, J. F., Barros, D., Castro, J. P. & Cammack, N. Fueling Open-Source Drug Discovery: 177 Small-Molecule Leads against Tuberculosis. *ChemMedChem* **8**, 313-321 (2013).
428 Hughes, J., Rees, S., Kalindjian, S. & Philpott, K. Principles of early drug discovery. *British Journal of Pharmacology* **162**, 1239-1249 (2011).
429 Özbal, C. C., LaMarr, W. A., Linton, J. R., Green, D. F., Katz, A., Morrison, T. B. & Brenan, C. J. H. High Throughput Screening via Mass Spectrometry: A Case Study Using Acetylcholinesterase. *ASSAY and Drug Development Technologies* **2**, 373-382 (2004).
430 Genick, C. C. & Wright, S. K. Biophysics: for HTS hit validation, chemical lead optimization, and beyond. *Expert Opinion on Drug Discovery* **12**, 897-907 (2017).
431 Wu, G., Yuan, Y. & Hodge, C. N. Determining Appropriate Substrate Conversion for Enzymatic Assays in High-Throughput Screening. *Journal of Biomolecular Screening* **8**, 694-700 (2003).
432 Analytical Methods, C. Robust statistics–how not to reject outliers. Part 1. Basic concepts. *Analyst* **114**, 1693-1697 (1989).
433 Coma, I., Herranz, J. & Martin, J. in *High Throughput Screening: Methods and Protocols, Second Edition* (eds William P. Janzen & Paul Bernasconi) 69-106 (Humana Press, 2009).
434 Baaske, P., Wienken, C. J., Reineck, P., Duhr, S. & Braun, D. Optical Thermophoresis for Quantifying the Buffer Dependence of Aptamer Binding. *Angewandte Chemie International Edition* **49**, 2238-2241 (2010).
435 Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular interaction studies using microscale thermophoresis. *Assay and drug development technologies* **9**, 342-353 (2011).
436 Stokes, S. S., Albert, R., Buurman, E. T., Andrews, B., Shapiro, A. B., Green, O. M., McKenzie, A. R. & Otterbein, L. R. Inhibitors of the acetyltransferase domain of N-acetylglucosamine-1-phosphate-uridylyltransferase/glucosamine-1-phosphate-acetyltransferase (GlmU). Part 2: Optimization of physical properties leading to antibacterial aryl sulfonamides. *Bioorganic & Medicinal Chemistry Letters* **22**, 7019-7023 (2012).
437 Pompeo, F., van Heijenoort, J. & Mengin-Lecreulx, D. Probing the Role of Cysteine Residues in Glucosamine-1-Phosphate Acetyltransferase Activity of the Bifunctional GlmU Protein fromEscherichia coli: Site-Directed Mutagenesis and Characterization of the Mutant Enzymes. *Journal of Bacteriology* **180**, 4799-4803 (1998).
438 Kelly, S. M. & Price, N. C. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Current Protein and Peptide Science* **1**, 349-384 (2000).

439 de Carvalho, L. P. S., Ling, Y., Shen, C., Warren, J. D. & Rhee, K. Y. On the chemical mechanism of succinic semialdehyde dehydrogenase (GabD1) from Mycobacterium tuberculosis. *Archives of Biochemistry and Biophysics* **509**, 90-99 (2011).

440 Alatossava, T., Jütte, H., Kuhn, A. & Kellenberger, E. Manipulation of intracellular magnesium content in polymyxin B nonapeptide-sensitized Escherichia coli by ionophore A23187. *Journal of Bacteriology* **162**, 413-419 (1985).

441 Argyrou, A. & Blanchard, J. S. Kinetic and Chemical Mechanism of Mycobacterium tuberculosis 1-Deoxy-d-xylulose-5-phosphate Isomeroreductase. *Biochemistry* **43**, 4375-4384 (2004).

442 Bellion, E. The Biological Chemistry of the Elements: The Inorganic Chemistry of Life (da Silva, J. J. R. Frausto; Williams, R. J. P.). *Journal of Chemical Education* **69**, A326 (1992).

443 D'Souza, V. M. & Holz, R. C. The methionyl aminopeptidase from Escherichia coli can function as an iron(II) enzyme. *Biochemistry* **38**, 11079-11085 (1999).

444 Craggs, P. D., Mouilleron, S., Rejzek, M., de Chiara, C., Young, R. J., Field, R. A., Argyrou, A. & de Carvalho, L. P. S. The Mechanism of Acetyl Transfer Catalyzed by Mycobacterium tuberculosis GlmU. *Biochemistry* **57**, 3387-3401 (2018).

445 Simms, S. A., Voige, W. H. & Gilvarg, C. Purification and characterization of succinyl-CoA: tetrahydrodipicolinate N-succinyltransferase from Escherichia coli. *Journal of Biological Chemistry* **259**, 2734-2741 (1984).

446 Hurtado-Guerrero, R., Raimi, O., Shepherd, S. & van Aalten, D. M. F. Glucose-6-phosphate as a probe for the glucosamine-6-phosphate N-acetyltransferase Michaelis complex. *FEBS Letters* **581**, 5597-5600 (2007).

447 Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L. & Woolfson, D. N. Carbohydrate–Aromatic Interactions in Proteins. *Journal of the American Chemical Society* **137**, 15152-15160 (2015).

448 Scheibner, K. A., De Angelis, J., Burley, S. K. & Cole, P. A. Investigation of the Roles of Catalytic Residues in SerotoninN-Acetyltransferase. *Journal of Biological Chemistry* **277**, 18118-18126 (2002).

449 Grimsley, G. R., Scholtz, J. M. & Pace, C. N. A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Science* **18**, 247-251 (2009).

450 Barbara Schowen, K. & Schowen, R. L. in *Methods in Enzymology* Vol. 87 (ed Daniel L. Purich) 551-606 (Academic Press, 1982).

451 Karsten, W. E., Lai, C.-J. & Cook, P. F. Inverse Solvent Isotope Effects in the NAD-Malic Enzyme Reaction Are the Result of the Viscosity Difference between D2O and H2O: Implications for Solvent Isotope Effect Studies. *Journal of the American Chemical Society* **117**, 5914-5918 (1995).

452 Venkatasubban, K. S. & Schowen, R. L. The Proton Inventory Techniqu. *Critical Reviews in Biochemistry* **17**, 1-44 (1984).

453 Gribbon, P., Lyons, R., Laflin, P., Bradley, J., Chambers, C., Williams, B. S., Keighley, W. & Sewing, A. Evaluating Real-Life High-Throughput Screening Data. *Journal of Biomolecular Screening* **10**, 99-107 (2005).

454 Harper, G. & Pickett, S. D. Methods for mining HTS data. *Drug Discovery Today* **11**, 694-699 (2006).

455 Zhu, T., Cao, S., Su, P.-C., Patel, R., Shah, D., Chokshi, H. B., Szukala, R., Johnson, M. E. & Hevener, K. E. Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *Journal of medicinal chemistry* **56**, 6560-6572 (2013).

456 Dahlin, J. L., Nissink, J. W. M., Strasser, J. M., Francis, S., Higgins, L., Zhou, H., Zhang, Z. & Walters, M. A. PAINS in the Assay: Chemical Mechanisms of Assay Interference and Promiscuous Enzymatic Inhibition Observed during a Sulfhydryl-Scavenging HTS. *Journal of Medicinal Chemistry* **58**, 2091-2113 (2015).

457 Nikolova, N. & Jaworska, J. Approaches to Measure Chemical Similarity – a Review. *QSAR & Combinatorial Science* **22**, 1006-1026 (2003).

458 Willett, P., Barnard, J. M. & Downs, G. M. Chemical Similarity Searching. *Journal of Chemical Information and Computer Sciences* **38**, 983-996 (1998).