The Biochemistry and Phylogeny of CysE: an antimicrobial target

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Molecular and Cellular Biology

at

The University of Waikato

by

Keely Elizabeth Ann Oldham

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2024

Abstract

Neisseria gonorrhoeae is an obligate human pathogen and the causative agent of the sexually transmitted infection (STI), gonorrhoea. There are over 87 million gonorrhoea cases per year globally, making it the second most prevalent bacterial STI. The rapid emergence of antimicrobial-resistant strains has led *N. gonorrhoeae* to being labelled as a priority pathogen for the development of new antimicrobials by the World Health Organisation. One strategy for identifying new antimicrobials is through targeting pathways required for virulence and mitigation of oxidative stress, such as sulphur assimilation pathways. This thesis investigates the potential of targeting the *de novo* cysteine biosynthetic enzyme, serine acetyltransferase, from *N. gonorrhoeae* (*Ng*CysE) for new antimicrobials.

In this thesis, we present the kinetic and structural characterisation of *Ng*CysE. *Ng*CysE has serine acetyltransferase activity and is sensitive to feedback inhibition by L-cysteine. Small-angle X-ray scattering demonstrates *Ng*CysE is a single hexameric species in solution with subtle conformational changes upon binding of its inhibitor L-cysteine. Using X-ray crystallography, we present two *Ng*CysE structures, with L-malate (2.01 Å) and substrate L-serine (2.8 Å), bound. *Ng*CysE crystallises as a homohexamer with 3:2-fold symmetry with each monomer containing a N-terminal α-helical SATase and C-terminal left-handed β-helix domain. Additionally, we see extensive density for the C-terminal α-helical tails because of domain swapping, resulting in one of the most complete CysE crystal structures to date.

The high resolution *Ng*CysE structure was used for structure-based virtual inhibitor screening to identify potential inhibitors of *Ng*CysE. Virtual screening produced 28 hit compounds that were tested *in vitro*, with the most potent inhibitor, compound 2, demonstrating inhibition in the low micromolar range. Analysis of compound 2 docking reveals interactions with both *Ng*CysE substrate binding sites, serine and acetyl-CoA. This is the first reported inhibitor of *Ng*CysE and provides a starting point for development of new antimicrobials for treating gonorrhoea infection.

Lastly, we present the phylogenetic analysis of truncated and full-length CysE isoforms across the bacterial kingdom. We demonstrate that a truncated isoform has

a discrete truncation of four α-helices from the N-terminal SATase domain and is the predominant isoform for Gram-positive bacteria. The functional implications of this truncation are explored using predictive modelling and demonstrate that CysE is likely to exist as a trimer rather than the hexamer seen for the full-length isoform. This highlights the diversity of CysE isoforms and paves the way for determining the evolutionary origin of these isoforms.

Collectively, this thesis highlights the merit in the emerging field of CysE inhibition. Moreover, we provide the basis for the future development of not only *Ng*CysE inhibitors, but also the development of CysE inhibitors for other bacterial pathogens.

.

Acknowledgements

Firstly, I would like to thank my Chief Supervisor Dr. Joanna Hicks. Jo, you have guided and supported me along my research journey every step of the way. Your generosity and support have given me countless opportunities to present my research, given me the confidence to take ownership of this project and explore new research avenues, and all together has helped me grow as a researcher. I would also like to thank my Co-supervisor Professor Vic Arcus for his guidance and support over the course of this thesis.

Secondly, a thank you to Dr. Wanting Jiao. Thank you for providing your expertise for this project and being a great collaborator.

A special thank you to Erica Prentice for her help in the analysis of the kinetic data and for overall being a great mentor and sanity checker. And a further thank you to Emma Summers for your structural expertise and analysis of the CysE structures.

Thank you Judith for being an amazing lab mum! You always go above and beyond - from ensuring all birthdays are celebrated (with willing or unwilling participants), to ensuring that the lab is clean and orderly (even if the person writing this is the lab hoarder). You do an amazing job and I will miss your lemon blueberry cake.

I would like to thank Emily for supporting me during this thesis. From navigating Hanmer Springs, to fixing AKTAs, to feeding me snacks, to even proofreading this thesis, you were the best lab buddy and I will miss our times in the lab.

Thank you Carlin for being my coding expert and scripting an amazing renaming tool.

I would like to thank Emma and Stacy for you being my gono support persons. And an extra thank you for Stacy for transforming the *cysE* compliment strain.

Additionally, thank you Meghan for proofreading this thesis.

Thank you to all the current and past students for making the lab a great place to work, from coffee at the lake, to extended morning teas; it made coming to the lab a joy.

I would like to thank The University of Waikato for funding me for the first three years of this thesis and for my supervisor Dr. Joanna Hicks for the months that followed.

Finally, thank you to my family for plying me with food, buying conference clothes, and always taking an interest in my work. Trust me, I'm a good investment!

Lastly, thank you Connor for your unwavering support; for being the son of an AKTA technician, an expert in 3D printing, and for also becoming an unwilling expert on this thesis topic. Now, onto the next adventure.

Table of Contents

List of Figures

List of Tables

List of Equations

Glossary

Acetyl-CoA	acetyl-coenzyme A
ABC	ATP-binding cassette
Abs	Absorbance
Arg	arginine, amino acid
Asp	aspartate, amino acid
ASU	asymmetric unit
bp	base pair
CoA	co-enzyme A
CDC	centre for disease and control prevention
CSC	cysteine synthase complex
CysE	serine acetyltransferase
DTNB	5,5′-Dithiobis(2-nitrobenzoic acid) / Ellman's reagent
DMSO	dimethyl sulphoxide
EDTA	ethylenediaminetetraacetic acid
FPLC	fast protein liquid chromatography
g	force of gravity
GCB	gonococcal base medium
Glu	glutamate, amino acid
Gly	glycine, amino acid
Hexahis tag	hexa-histidine tag
His	histidine, amino acid
HMM	hidden Markov model
IC_{50}	inhibitor concentration where activity is reduced by 50%
IMAC	immobilised metal affinity chromatography
IPTG	isopropylthio-β-galactosidase
K_{av}	gel phase distribution co-efficient
k_{cat}	enzyme rate constant
k_{cat}/K_M	catalytic efficiency constant
K_d	dissociation constant
kDa	kilo Dalton
K_i	inhibitor constant
K_M	Michaelis constant
KO	knockout
LB	Luria Bertani
L-cys	L-cysteine
L-ser	L-serine
LβH	Left-handed parallel β-helix
MD	molecular dynamics
MM-GBSA	molecular mechanics-generalised Born surface area
MMK	Michaelis Menten kinetics
MR	molecular replacement
MWCO	molecular weight cut off
NAS	*N*-acetylserine
OAS	*O*-acetylserine

OASS-A	*O*-acetylserine sulphydrylase A (CysK)
OASS-B	*O*-acetylserine sulphydrylase B (CysM)
PAGE	polyacrylamide gel electrophoresis
PAPS	3’-phosphoadenosine-5’-phosphosulphate
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
pI	isoelectric point
pLDDT	predicted local distance difference test
PLP	pyridoxal 5’-phosphate
PMN	polymorphonuclear neutrophil
Pro	proline, amino acid
QM/MM	quantum mechanics/ molecular mechanics
R*g*	radius of gyration
RT	room temperature
s	second(s)
SAT	serine *O*-acetyltransferase
SATase	serine *O*-acetyltransferase
SDS	sodium dodecyl sulphate
SEM	standard error of mean
Ser	serine, amino acid
SSN	sequence similarity network
STI	sexually transmitted infection
tNCS	translational non-crystallographic symmetry
TEMED	tetramethylethylenediamine
TNB	2-nitro-5-thiobenzoate
Tris	tris(hydroxymethyl)aminomethane
V	volts
Val	valine, amino acid
V_{max}	maximum reaction rate
WHO	world health organisation
XDS	X-ray detector software

Chapter One: Introduction

1.1 Introduction

Neisseria gonorrhoeae is the causative agent of the sexually transmitted infection (STI), gonorrhoea. This obligate human pathogen is the second most prevalent bacterial STI, with 87 million new infections annually (Rowley et al., 2019). Since the discovery of antibiotics, *N. gonorrhoeae* has successfully developed resistance to all antibiotic classes used in gonorrhoea treatment (Unemo & Shafer, 2011), including the frontline antibiotic, ceftriaxone (Day et al., 2022; Eyre et al., 2018; Jennison et al., 2019). The increasing prevalence of ceftriaxone-resistant strains, together with rapidly increasing incidence rates, has led to *N. gonorrhoeae* being labelled by the World Health Organisation as a priority pathogen for the development of new antimicrobials (WHO, 2017). Current antibiotics used in gonorrhoea treatment target essential bacterial processes, such as cell wall synthesis and protein translation, but there is a demand for the identification of novel antimicrobial targets to combat antibiotic resistance. One strategy is to target pathways that contribute to pathogen persistence and virulence, such as sulphur metabolism. We propose that the *de novo* biosynthesis pathway for the sulphur containing amino acid L-cysteine, is an ideal candidate for antimicrobial development for *N. gonorrhoeae*, specifically targeting the rate-limiting enzyme, serine acetyltransferase.

De novo cysteine biosynthesis is well-conserved across both bacteria and higher plants. *De novo* synthesis is the primary mechanism that allows bacteria to acquire and assimilate inorganic sulphur sources, sulphate and thiosulphate, incorporating them into the amino acid, L-cysteine (Kredich, 2008). L-cysteine is a versatile thiol donor and feeds into synthesis of key thiol metabolites, such as co-enzyme A, iron sulphur clusters, L-methionine and key reducing agents, such as glutathione. The synthesis of cysteine is a dual-step process; the reaction begins with the acetylation of L-serine by serine acetyltransferase (CysE/SAT) producing *O*-acetylserine (Kredich & Tomkins, 1966), which is then condensed with sulphide by *O*-acetylserine sulphydrylase-A (CysK/OASS-A) (Kredich & Tomkins, 1966) or with thiosulphate by alternate isoform *O*-acetylserine sulphydrylase-B (CysM/OASS-B) (Becker & Tomkins, 1969), to produce L-cysteine (Kredich, 2008). This pathway

is regulated through classical feedback inhibition via L-cysteine binding to CysE (Benoni et al., 2017; Kredich & Tomkins, 1966). Recently, there has been increased research into the role of cysteine biosynthesis pathways in bacteria, with a particular focus on how inhibition of this pathway could be exploited for the development of novel antimicrobials and antibiotic enhancers (Campanini et al., 2015; Hicks, Oldham, McGarvie, & Walker, 2022). In this thesis, we explore how targeting *de novo* cysteine biosynthesis pathway is a promising avenue for developing new antimicrobials against bacteria with a focus on the human pathogen *N. gonorrhoeae*.

Currently, the literature classifies *N. gonorrhoeae* as a cysteine auxotroph (Catlin, 1973), however there is evidence that supports the ability of *N. gonorrhoeae* to grow on alternate sulphur sources in the absence of cysteine (Le Faou, 1984). The investigation of sulphur acquisition and assimilation in *N. gonorrhoeae* is being investigated by my Chief Supervisor Dr. Joanna Hicks and members of our research group. Previous work by our group extensively reviewed and analysed the *de novo* cysteine biosynthetic pathway in *Neisseria* species (Hicks & Mullholland, 2018). Genomic interrogation revealed that the sulphate reduction arm of the *de novo* cysteine biosynthesis pathway was present in *N. meningitidis*, but a large 3.5 kB genomic deletion in *N. gonorrhoeae* disrupts this sulphate reduction pathway, rendering *N. gonorrhoeae* incapable of using sulphate as a sulphur donor. Growth experiments conducted by Le Faou et al. (1984), demonstrated that *N. gonorrhoeae* could not utilise sulphate, but instead could use thiosulphate as a sole sulphur source (Le Faou, 1984). The cysteine biosynthetic enzymes that remain genomically intact in *N. gonorrhoeae* are a sulphate/thiosulphate ABC importer (*cysA*, *cysW*, *cysT* and *sbp*), and the key enzymes serine acetyltransferase (*Ng*CysE) and *O*-acetylserine sulphydrylase-A sulphydrylase (*Ng*CysK). Given the disruption of the sulphate reduction arm, questions were raised over the functionality of the remaining *Ng*CysE and *Ng*CysK. Preliminary characterisation conducted as a part of my Master's studies demonstrated that *Ng*CysE has serine acetyltransferase activity comparable to other bacterial homologues (Oldham, 2020). Additionally, characterisation of *Ng*CysK demonstrated that CysK only showed activity with hydrogen sulphide and is incapable of using thiosulphate as a sulphur donor (McGarvie, 2021). Based on the observation that *Ng*CysE and *Ng*CysK are functional enzymes, an alternative sulphurtransferase-dependent pathway to obtain

sulphide from thiosulphate and glutathione, that has been characterised in *E. coli* (Kawano et al., 2017), is being investigated in *N. gonorrhoeae*, by PhD student Stacy van Niekerk. Initial characterisation of this pathway (van Niekerk, 2021) demonstrated that the sulphurtransferases present in *N. gonorrhoeae* were able to produce sulphide from thiosulphate. Overall, on the basis of our findings, we propose that *N. gonorrhoeae* is capable of *de novo* cysteine biosynthesis.

This research focuses on elucidating the role of serine acetyltransferases in bacteria and how this can be explored for antimicrobial development. This thesis follows on from preliminary characterisation work completed during my Master's studies. The work presented in this thesis details the first *in vitro* kinetic and structural characterisation of *Ng*CysE. *Ng*CysE crystal structures were used for structure-based virtual inhibitor screening to identify novel *Ng*CysE inhibitors. Additionally, this thesis provides an in-depth review of antimicrobial targeting efforts of the *de novo* cystine biosynthetic pathway in bacterial pathogens. Finally, we have investigated the diversity and distribution of full-length and truncated CysE isoforms across bacterial kingdom and investigate the implications of this truncation on CysE function.

This thesis contains published articles, submitted manuscripts, and manuscripts prepared for submission. These articles are presented in thesis are a literature review and three manuscripts, followed by a Conclusion and Future Directions Chapter. The literature review is presented in Chapter Two. The citation for the review is as follows:

Hicks, J. L., **Oldham, K. E. A**., McGarvie, J., & Walker,. E. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. *Biosci Reports*, *42*(10) doi: 10.1042/bsr20220368

1.2 Thesis Objectives

The overall aims of this thesis are to firstly; elucidate the role of serine acetyltransferase in the human pathogen *N. gonorrhoeae* and secondly, to improve our understanding of the diversity and phylogeny of serine acetyltransferase enzymes in bacteria. To achieve this, the thesis objectives are as follows:

Objective One:

- To biochemically and structurally characterise serine acetyltransferase from *N. gonorrhoeae*.

This research is presented in Chapter Three as a published peer-reviewed article:

- **Oldham, K. E. A**., Prentice, E. J., Summers, E. L., & Hicks, J. L. (2022). Serine acetyltransferase from *Neisseria gonorrhoeae*; structural and biochemical basis of inhibition. *The Biochemical journal*, *479*(1), 57-74. doi:10.1042/bcj20210564

Objective Two:

- To identify and characterise novel CysE inhibitors targeting serine acetyltransferase from *N. gonorrhoeae*.

This research is presented in Chapter Four as a manuscript prepared for submission:

Oldham, K. E. A., Jiao, W., & Hicks, J. L. (2023). Identification of novel inhibitors targeting serine acetyltransferase from *Neisseria gonorrhoeae.* *(ready for submission)*.

Objective Three:

- Investigate the phylogenetic diversity and distribution of bacterial serine acetyltransferases.

This research is presented in Chapter Five as a manuscript prepared for submission:

Oldham, K. E. A., & Hicks, J. L. (2024). Curious case of CysE: Distribution and diversity of serine acetyltransferase isoforms. *(ready for submission)*

1.3 Thesis Outline

This thesis consists of six chapters titled; Introduction, Literature Review, Characterisation of serine acetyltransferase from *Neisseria gonorrhoeae*, Identification and Characterisation of Inhibitors of serine acetyltransferase from *Neisseria gonorrhoeae*, and Diversity and Distribution of CysE isoforms. In several chapters, additional future direction sections have been included and are highlighted in blue text.

1.4 Acknowledgement of Funds

I was funded by a Doctoral Scholarship from the University of Waikato. This research was funded by grants from the Waikato Medical Research Foundation and the Health Research Council (Grant No. 19/602) awarded to Chief Supervisor Dr. Joanna Hicks. Virtual inhibitor screening was funded by the Maurice Wilkins Centre. Access to the Australian Synchrotron macromolecular crystallography (MX) and small-angle X-ray scattering (SAXS) beamlines were funded through the New Zealand Synchrotron Group.

1.5 References

Becker, M. A., & Tomkins, G. M. (1969). Pleiotropy in a Cysteine-requiring Mutant of Salmonella typhimurium Resulting from Altered Protein-Protein Interaction. *Journal of Biological Chemistry, 244*(21), 6023-6030. 10.1016/S00219258(18)63576-8

Benoni, R., De Bei, O., Paredi, G., Hayes, C. S., Franko, N., Mozzarelli, A., . . . Campanini, B. (2017). Modulation of Escherichia coli serine acetyltransferase catalytic activity in the cysteine synthase complex. *FEBS letters, 591*(9), 1212-1224. 10.1002/18733468.12630

Campanini, B., Pieroni, M., Raboni, S., Bettati, S., Benoni, R., Pecchini, C., . . . Mozzarelli, A. (2015). Inhibitors of the Sulphur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Current Medicinal Chemistry, 22*(2), 187-213. 10.2174/0929867321666141112122553

Catlin, B. W. (1973). Nutritional Profiles of Neisseria gonorrhoeae, Neisseria meningitidis, and Neisseria lactamica in Chemically Defined Media and the Use of Growth Requirements for Gonococcal Typing. *The Journal of Infectious Diseases, 128*(2), 178-194. 10.1093/infdis/128.2.178

Day, M., Pitt, R., Mody, N., Saunders, J., Rai, R., Nori, A., . . . Fifer, H. (2022). Detection of 10 cases of ceftriaxone-resistant Neisseria gonorrhoeae in the United Kingdom, December 2021 to June 2022. *Euro Surveill, 27*(46) 10.2807/15607917.Es.2022.27.46.2200803

Eyre, D. W., Sanderson, N. D., Lord, E., Regisford-Reimmer, N., Chau, K., Barker, L., . . . Andersson, M. I. (2018). Gonorrhoea treatment failure caused by a Neisseria gonorrhoeae strain with combined ceftriaxone and high-level azithromycin resistance, England, February 2018. *Euro surveillance : bulletin*

Europeen sur les maladies transmissibles = European communicable disease bulletin, 23(27), 1800323. 10.2807/1560-7917.ES.2018.23.27.1800323

Hicks, J. L., & Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology, 164*(12), 1471-1480. 10.1099/mic.0.000728

Hicks, J. L., Oldham, K. E. A., McGarvie, J., & Walker, E. J. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. *Biosci Rep, 42*(10) 10.1042/bsr20220368

Jennison, A. V., Whiley, D., Lahra, M. M., Graham, R. M., Cole, M. J., Hughes, G., . . . Eyre, D. (2019). Genetic relatedness of ceftriaxone-resistant and high-level azithromycin resistant Neisseria gonorrhoeae cases, United Kingdom and Australia, February to April 2018. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 24*(8), 1900118. 10.2807/1560-7917.ES.2019.24.8.1900118

Kawano, Y., Onishi, F., Shiroyama, M., Miura, M., Tanaka, N., Oshiro, S., . . . Ohtsu, I. (2017). Improved fermentative L-cysteine overproduction by enhancing a newly identified thiosulphate assimilation pathway in Escherichia coli. *Appl Microbiol Biotechnol, 101*(18), 6879-6889. 10.1007/s00253-017-8420-4

Kredich, N. (2008). Biosynthesis of Cysteine. *EcoSal Plus* doi:10.1128/ecosalplus.3.6.1.11

Kredich, N. M., & Tomkins, G. M. (1966). The Enzymic Synthesis of l-Cysteine in Escherichia coli and Salmonella typhimurium. *Journal of Biological Chemistry, 241*(21), 4955-4965.

Le Faou, A. (1984). Sulphur nutrition and metabolism in various species of Neisseria. *Annales de l'Institut Pasteur / Microbiologie, 135*(1, Supplement B), 3-11. 10.1016/S0769-2609(84)80037-X

McGarvie, J. (2021). *Structural and biochemical characterisation of O-acetylserine sulphydrylase (CysK) from Neisseria gonorrhoeae* (Masters). The University of Waikato, Hamilton, New Zealand. Retrieved from https://hdl.handle.net/10289/14596

Oldham, K. (2020). *Cysteine biosynthesis and the role of CysE in Neisseria gonorrhoeae* (Masters). The University of Waikato, Hamilton, New Zealand. Retrieved from https://hdl.handle.net/10289/13461

Rowley, J., Vander Hoorn, S., Korenromp, E., Low, N., Unemo, M., Abu-Raddad, L. J., . . . Taylor, M. M. (2019). Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bulletin of the World Health Organisation, 97*(8), 548-562P. 10.2471/BLT.18.228486

Unemo, M., & Shafer, W. M. (2011). Antibiotic resistance in Neisseria gonorrhoeae: origin, evolution, and lessons learned for the future. *Ann N Y Acad Sci, 1230*, E19-28. 10.1111/j.1749-6632.2011.06215.x van Niekerk, S. L. (2021). *Alternative sulphur acquisition pathways in Neisseria gonorrhoeae* (Masters). The University of Waikato, Hamilton, New Zealand. Retrieved from https://hdl.handle.net/10289/14597

WHO. (2017). World Health Organisation (WHO) Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics. Geneva: WHO.

Chapter Two: Literature Review

2.1 Preface

This chapter provides an in-depth examination of current research into the development of antimicrobials and adjuvants targeting the *de novo* cysteine biosynthetic enzymes serine *O*-acetyltransferase (CysE) and *O*-acetylserine sulphydrylase (CysK/OASS-A; CysM/OASS-B). This review provides contextual merit for antimicrobial targeting of the *de novo* cysteine biosynthesis, through highlighting the crucial role cysteine biosynthesis plays in bacterial pathogenesis, particularly in bacterial persistence and antibiotic resistance, and how the enzymes CysE and CysK/CysM are implicated. Here, we have focused on the extensive characterisation of CysE and the CysK/CysM bacterial isoforms and current strategies employed for identification and optimisation of new inhibitors and how this is being explored for antimicrobials and adjuvants against bacteria, including antimicrobial resistant strains.

This review was commissioned based on the publication of our research article characterising CysE from *Neisseria gonorrhoeae* (Chapter 3) and was published in *Bioscience reports* (Portland Press) as a peer-reviewed Review Article. To keep the review up to date with current research, new information is included in this chapter and is formatted as blue text to differentiate it from the publication text. This text is present in the CysE inhibitor section (Section 2.5.6) The work presented in this chapter is formatted as a published review. Citation for this publication is as follows:

Hicks, J. L., **Oldham, K. E. A.,** McGarvie, J., & Walker, E. J. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. *Bioscience reports*, *42*(10), BSR20220368. 10.1042/BSR20220368.

2.2 Author contributions

Given that our research group are investigating both CysE and CysK enzymes from *N. gonorrhoeae,* the collation of literature and manuscript writing was conducted collaboratively between myself and co-authors Jack McGarvie, Emma Walker and my Chief Supervisor, Joanna Hicks. I wrote the sections pertaining to CysE and their inhibitors (section 1.4) and created figures for these sections and the

introduction. Jack McGarvie wrote the CysK section, and the CysK inhibitor section was written by both Emma Walker and Jack McGarvie. Joanna Hicks wrote the introduction and conclusion. All authors contributed equally to editing of the manuscript. Co-author contributions can be found in Appendix D.

Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens.

Joanna Hicks[1], Keely Oldham[1,2], Jack McGarvie[1], Emma Walker[2]

[1]*Te Huataki Waiora School of Health, University of Waikato, New Zealand*

[2]*Te Aka Mātuatua School of Science, University of Waikato, New Zealand*

2.3 Abstract

Antibiotics are the cornerstone of modern medicine and agriculture, and rising antibiotic resistance is one the biggest threats to global health and food security. Identifying new and different druggable targets for the development of new antibiotics is absolutely crucial to overcome resistance. Adjuvant strategies that either enhance the activity of existing antibiotics or improve clearance by the host immune system provide another mechanism to combat antibiotic resistance. Targeting a combination of essential and non-essential enzymes that play key roles in bacterial metabolism is a promising strategy to develop new antimicrobials and adjuvants, respectively. The enzymatic synthesis of L-cysteine is one such strategy. Cysteine plays a key role in proteins and is crucial for the synthesis of many biomolecules important for defence against the host immune system. Cysteine synthesis is a two-step process, catalysed by two enzymes. Serine acetyltransferase (CysE) catalyses the first step to synthesise the pathway intermediate *O*-acetylserine, and *O*-acetylserine sulphydrylase (CysK/CysM) catalyses the second step using sulphide or thiosulphate to produce cysteine. Disruption of the cysteine biosynthesis pathway results in dysregulated sulphur metabolism, altering the redox state of the cell leading to decreased fitness, enhanced susceptibility to oxidative stress and increased sensitivity to antibiotics. In this review, we summarise the structure and mechanism of characterised CysE and CysK/CysM enzymes from a variety of bacterial pathogens, and the evidence that support targeting these enzymes for the development of new antimicrobials or antibiotic adjuvants. In addition, we explore and compare compounds identified thus far that target these enzymes.

2.4 Introduction

Antibiotic resistance is a slow burning global pandemic that threatens not just human health and life expectancy but also food production. The discovery of antibiotics in the 1920s and 30s was a game changer for human health and agriculture. These drugs have saved millions of lives from previously fatal infections and massively reduced risk from surgical interventions. Within agriculture, antibiotics have been used for animal health and, in healthy food producing animals to promote growth and prevent disease. Yet not 100 years after the 'golden era' of antibiotic discovery we are facing the problem of extensively drug resistant (XDR) and multidrug resistant (MDR) strains of bacterial pathogens. The rising emergence of antibiotic resistance and the lack of new antibiotic classes discovered over the past 60 years requires new strategies to overcome resistance and target bacterial pathogens. Identifying novel druggable targets different from those currently targeted by antimicrobials are crucial to overcoming antibiotic resistance. Along with this, adjuvant strategies targeting nonessential targets such as those important for bacterial virulence, persistence or host colonisation are gaining interest (Becker et al., 2006; Bhave, Muse, & Carroll, 2007). Targeting these enzymes that are nonessential during the normal bacterial life cycle but become essential during infection and host invasion, could decrease the incidence of antimicrobial resistance (AMR), as inhibition of these nonessential targets would facilitate clearance by the immune system without stimulating resistance. Targeting nonessential enzymes often decreases bacterial fitness, thereby inhibitors of these enzymes could act as adjuvants to enhance the potency of existing antibiotics. Most pathogens spend at least part of their life cycle in an extremely challenging environment; infection and survival within the hostile host environment relies on a series of sulphur-containing molecules, including Fe-S clusters, thiamine, thioredoxin, glutathione and biotin, which have detoxifying capabilities and reducing power (Fasnacht & Polacek, 2021; Roop, Gaines, Anderson, Caswell, & Martin, 2009).

Cysteine is absolutely crucial for the synthesis of sulphur-containing biomolecules; therefore, inhibiting cysteine synthesis is a promising strategy for both potential new antimicrobials and antimicrobial adjuvants. Inhibition of cysteine biosynthesis has been proven to interfere with a pathogen's ability to fight oxidative stress, infect

the host and establish long-term infection (Abd El-Aleam, George, Georgey, & Abdel-Rahman, 2021; Brunner et al., 2017; Campanini, Pieroni, et al., 2015). For example, cysteine metabolism is a promising drug target in *Salmonella enterica* serovar Typhimurium (Becker & Tomkins, 1969; Turnbull & Surette, 2008, 2010) and *Mycobacterium tuberculosis* (Brunner et al., 2016; Brunner et al., 2017; Burns-Huang & Mundhra, 2019) where suppression or reduction of cysteine synthesis led to decreased fitness and infectivity. Inhibition of cysteine biosynthesis has also been associated with a dysregulated oxidative stress response, enhancing the antimicrobial activity of existing antibiotics (Campanini, Pieroni, et al., 2015; Turnbull & Surette, 2008, 2010). Mammals lack the biosynthetic machinery for the *de novo* synthesis of cysteine from inorganic sulphur, relying on the reverse transulphuration of dietary methionine to obtain cysteine. Whereas bacteria and plants have highly conserved enzymes for the assimilation of inorganic sulphur into cysteine (Stipanuk, Dominy, Lee, & Coloso, 2006).

The synthesis of cysteine is a two-step process, catalysed by two enzymes, serine acetyltransferase (SAT;CysE) catalyses the first step, requiring L-serine and acetyl-coenzyme A (acetyl-CoA) to produce *O*-acetylserine (OAS). *O*-acetylserine sulphydrylase A (OASS-A, CysK) combines *O*-acetylserine with sulphide to produce cysteine, whereas the OASS-B (CysM) isoform can utilise both sulphide and thiosulphate as the sulphur source for the synthesis of cysteine. The reductive assimilation of sulphate to sulphide for the synthesis of cysteine is catalysed by a suite of enzymes, often found in the sulphate reduction operon in bacteria (Figure 2.1). Thiosulphate is an alternative sulphur source used directly to synthesise cysteine by the OASS-B isoform, CysM (Figure 2.1) The enzymes required for sulphate assimilation are also being explored as potential antimicrobial targets but are outside the scope of this review.

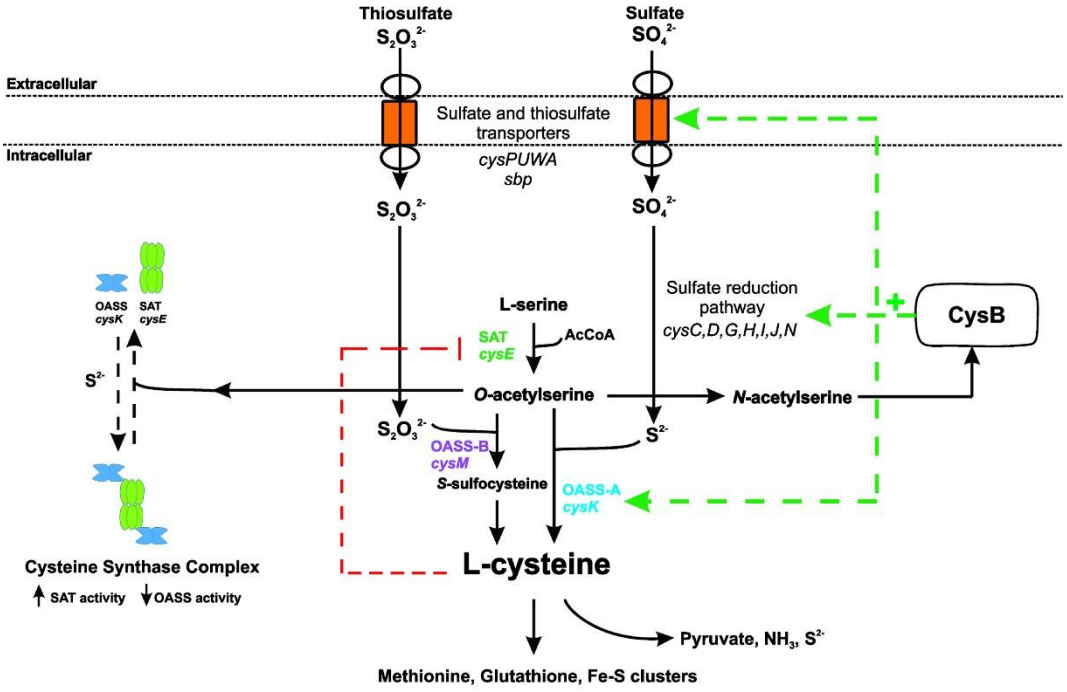


Figure 2.1 Overview of sulphur acquisition and assimilation pathways in bacteria. Inorganic sulphur uptake and assimilation pathways converge to be separately condensed with *O*-acetylserine to form cysteine. Transcriptional regulator, CysB, up-regulates the sulphur uptake (*cysPUWA*, *sbp*) and sulphur reduction pathways (*cysC,D,G,H,I,J,N*), in the presence of the readily forming *O*-acetylserine isoform, *N*-acetylserine. Cysteine feeds into *de novo* synthesis of sulphur-containing metabolites, such as methionine, glutathione, thioredoxin and Fe-S clusters. Cysteine inhibits CysE via feedback inhibition. Cysteine can be catabolised by cysteine desulfhydrase to release pyruvate, ammonia and sulphide.

The pathways of sulphur acquisition converge at cysteine (Figure 2.1). Therefore, the regulation of cysteine synthesis acts to control sulphur flux within the cell. Regulation occurs at both the genetic and protein levels. At the genetic level, the transcription factor CysB, belonging to the LysR family of transcriptional regulators, controls expression of key transporters of sulphur containing molecules, the sulphate reduction operon, and enzymes involved in cysteine synthesis, with the exception of CysE, which is not regulated by CysB (Figure 2.1) (Colyer & Kredich, 1996; Guédon & Martin-Verstraete, 2007; Hryniewicz & Kredich, 1994; Kredich, 1992; Lochowska, Iwanicka-Nowicka, Plochocka, & Hryniewicz, 2001). Whereas at the protein level, CysE which catalyses the first step in the two-step reaction is inhibited by L-cysteine (Kredich & Tomkins, 1966) and also forms a complex with CysK, termed the cysteine synthase complex (CSC) (Figure 2.1). While part of the CSC, CysE activity is enhanced (Benoni, De Bei, et al., 2017), whereas CysK activity is inhibited due to the C-terminal peptide of CysE binding and occluding

the active site of CysK (Benoni, De Bei, et al., 2017). Pathway intermediate *O*-acetylserine and its isomer *N*-acetylserine, along with sulphide, regulate formation of the CSC and act as inducers and anti-inducers of CysB, respectively.

In this review, we briefly summarise the structural and mechanistic features of the CysE and CysK/CysM enzymes from bacterial pathogens and the evidence that support targeting this pathway for the development of new antimicrobials. We provide a comparison of compounds identified thus far that inhibit the CysE and OASS enzymes and the methods used to identify these compounds.

2.5 Serine acetyltransferase (CysE)

Serine acetyltransferase (CysE) CysE is a serine acetyltransferase that catalyses the first committed step of the cysteine biosynthetic pathway (Figure 2.1), utilising L-serine and acetyl-CoA to synthesize the pathway intermediate *O*-acetylserine. Not only does CysE catalyse the first committed step, it is subject to feedback inhibition by the pathway end-product L-cysteine. CysE is nonessential in some bacterial pathogens, but curiously is essential in others, suggesting that CysE inhibition holds promise as a new antimicrobial target and/or as an antibacterial adjuvant.

2.5.1 Essentiality and role of CysE during infection

CysE is important not only for the *de novo* synthesis of cysteine but also plays a key role in bacterial virulence. CysE is essential in the pathogens, *Staphylococcus aureus* (Chaudhuri et al., 2009), *Escherichia coli* O157:H7 strain (Warr et al., 2019), *Haemophilus influenzae* (Akerley et al., 2002) and the pathogenic *Neisseria* species; *Neisseria gonorrhoeae* (Remmele et al., 2014) and *Neisseria meningitidis* (Muir et al., 2020). Essentiality of *cysE* in these bacteria was elucidated using transposon mutagenesis screens, and requires further validation. Interestingly, all screens were performed in culture media containing cysteine and other organic sulphur compounds, indicating that these bacteria have a requirement for the *de novo* synthesis of cysteine, despite the availability of extracellular organic sulphur sources. Furthermore, some of these organisms have non-functional sulphate assimilation pathways, such as *N. gonorrhoeae* (Hicks & Mullholland, 2018) and *S. aureus* (Lithgow, Hayhurst, Cohen, Aharonowitz, & Foster, 2004), precluding

the reduction of sulphate to sulphide as a source of sulphur for cysteine synthesis. While these organisms can grow on the alternative sulphur sources, thiosulphate and sulphide (Le Faou, 1984; Lithgow et al., 2004), the essentiality of *cysE* in cysteine rich media suggests capability for the *de novo* synthesis of cysteine. It is also possible that CysE and/or the product *O*-acetylserine has an as yet unidentified function that makes it essential in these organisms.

In bacteria where CysE is nonessential and able to be deleted from the bacterial chromosome, growth defects and reduced virulence have been observed. For example, in the drug-resistant pathogen *Klebsiella pneumoniae*, CysE is not essential, but when deleted the mutant exhibits decreased fitness in a mouse model of pneumonia, thereby playing an important role in lung infection (Bachman et al., 2015). Another key example is *M. tuberculosis,* where randomised transposon mutagenesis studies demonstrated profound effects for a number of sulphate reduction genes and *cysE* by gene disruption (Rengarajan, Bloom, & Rubin, 2005). Further investigation demonstrated attenuation of *M. tuberculosis cysE* deletion strains in *in vitro* models of dormancy (Sassetti & Rubin, 2003).

Furthermore, *cysE* deletion in the sheep pathogen *Brucella ovis* resulted in poor growth in rich media and an early entry into stationary phase (Varesio, Fiebig, & Crosson, 2021). *B. ovis* Δ*cysE* strains were more susceptible to oxidative stress, shown through increased sensitivity to hydrogen peroxide. Cell invasion assays revealed that deletion of *cysE* did not affect cell infection but did significantly reduce replication within macrophages (Varesio et al., 2021). While the deletion of *cysE* is non-lethal, it imposes a fitness cost on *B. ovis* during intracellular growth, demonstrating the requirement for cysteine biosynthesis for survival within the host.

Deletion of *cysE* can also influence bacterial antibiotic resistance. For example, an *E. coli* K12 *cysE* deletion strain had increased tolerance to the antibiotic novobiocin (Rakonjac, Milic, & Savic, 1991). Conversely in the pathogen *S. typhimurium* loss of CysE function increased mecillinam sensitivity (Oppezzo & Antón, 1995). Supporting this increased antibiotic sensitivity phenotype, cysteine biosynthesis is crucial for swarm cell differentiation in *S. typhimurium*. Inactivation of cysteine biosynthetic genes resulting in cysteine auxotrophy led to increased antibiotic sensitivity in the swarm cell state (Turnbull & Surette, 2008), even though the

swarm medium contained sufficient cysteine to support growth. There is a complex interplay between cysteine metabolism, oxidative stress and antibiotic resistance, under normal growth conditions. *S. typhimurium* cysteine auxotrophs are oxidatively stressed and supporting this, in wild-type cells oxidative stress induces cysteine biosynthesis. *S. typhimurium cysE* deletion strains are incapable of synthesising cysteine and have decreased concentrations of reduced thiols or decreased total glutathione, leading to increased susceptibility to oxidative stress (Turnbull & Surette, 2010). Differences seen between *E. coli* with increased tolerance to novobiocin and *S. typhimurium* could be due to the mechanism of action of the antibiotic or the presence/absence of cysteine/cystine or inorganic sulphur compounds in the culture media used in experiments.

CysE from *E. coli* and *Providencia stuartii* impacts biofilm formation in these pathogens (Sturgill, Toutain, Komperda, O'Toole, & Rather, 2004). Deletion of *cysE* from *E. coli* and *P stuartii* enhanced biofilm formation. However, this could be reversed by supplementation with cysteine (100 μM) or *O*-acetylserine (10 mM) but not *N*-acetylserine (Sturgill et al., 2004). The high concentration of *O*-acetylserine required for biofilm reduction compared to cysteine, suggests cysteine itself negatively regulates biofilm formation in a yet to be determined role. Given that bacteria in biofilms are less sensitive to antibiotics, inhibiting their formation could provide a novel way for enhancing current antibiotics.

2.5.2 Structural Characteristics of CysE

CysE (EC 2.3.1.30) belongs to the acetyltransferase family of hexapeptide acyltransferases. Members of this family are defined by a six-peptide tandem repeat, [LIV]-[GAED]-X_2-[STAV]-X, which gives rise to a distinctive left-handed beta helix (LβH) (Raetz & Roderick, 1995). Structural characterisation of CysE enzymes with and without substrates and cysteine (inhibitor) bound from a range of Gram-negative bacterial pathogens including *E. coli* (1T3D)(Pye, Tingey, Robson, & Moody, 2004), *H. influenzae* (1SSM, 1SSQ, 1SST) (Olsen, B., Vetting, & Roderick, 2004), *Yersinia pestis* (3GVD) (Kim, Zhou, Peterson, Anderson, & Joachimiak, 2006), *Brucella melitensis* (3MC4), *Vibrio cholerae* (4H7O), *Brucella abortus* (4HZC, 4HZD) (Kumar, Kumar, Alam, & Gourinath, 2014), *K. pneumoniae* (6JVU) (Verma et al., 2020), *N. gonorrhoeae* (6WYE, 7RA4)

(Oldham, Prentice, Summers, & Hicks, 2022) and *S. typhimurium* (7E3Y). These structures provide insight into active site architecture upon substrate and inhibitor binding, which can be used to inform inhibitor design.

The CysE monomer consists of an amino-terminus (N-terminus) α- helical domain and a carboxy terminus (C-terminus) LβH domain (Figure 2.2A). The monomers assemble to form a trimer, which in turn forms a functional hexamer through hydrophobic trimer–trimer interactions via the alpha helices of the N-terminal domains (Figure 2.2A) (Gorman & Shapiro, 2004; Oldham et al., 2022; Olsen et al., 2004; Pye et al., 2004). There are six active sites in the hexamer, formed between adjacent monomers of the C-terminal LβH domain. There is one deviation from the hexapeptide repeat, producing a meandering loop which forms part of the active site (Pye et al., 2004).

The hexameric structure of CysE enzymes differs from other members of the acetyltransferase family, which are active trimers. The CysE hexamer forms the cysteine synthase complex (CSC) with the OASS- A/CysK enzyme. There are exceptions to this hexameric configuration, such as the CysE isoforms from the protozoan parasite *Entamoeba histolytica*, where CysE is an active trimer (Kumar et al., 2014).

2.5.3 CysE enzyme mechanism

Tight control of intracellular cysteine levels is essential for meeting the cysteine requirements of the cell, while preventing unwanted toxic effects of high cysteine concentrations (Park & Imlay, 2003). As mentioned previously, CysE is constitutively expressed (Noji et al., 2001), where the main form of regulation is through formation of the CSC and feedback inhibition by the pathway product L-cysteine (Kredich & Tomkins, 1966). L-cysteine is a potent inhibitor of bacterial

CysE enzymes with IC_{50} values of 0.5–10 μM reported (Benoni, Beck, et al., 2017; Hindson & Shaw, 2003; Kredich & Tomkins, 1966; Oldham et al., 2022). Kinetic studies show that cysteine is a competitive inhibitor relative to serine, through binding to the serine binding pocket which has been confirmed through crystallography (Kumar et al., 2014; Olsen et al., 2004; Pye et al., 2004) (Figure 2.2). Interestingly, cysteine displays competitive inhibition relative to acetyl-CoA

even though it binds to the serine binding pocket (Hindson, 2003; Johnson, Huang, Roderick, & Cook, 2004b; Oldham et al., 2022). This competitive inhibition is explained by observing CysE crystal structures with L-cysteine bound. Upon binding of L-cysteine in the serine binding site the C-terminal tail folds up against the CysE monomer, physically blocking the active site and preventing the binding of acetyl-CoA (Olsen et al., 2004). Supporting this, truncation of the last ten C-terminal residues that form the C-terminal tail reduces CysE sensitivity to cysteine inhibition (Mino, Yamanoue, et al., 2000). This is thought to prevent the accidental acetylation of L-cysteine, given its structural similarity to serine (Mino, Yamanoue, et al., 2000).

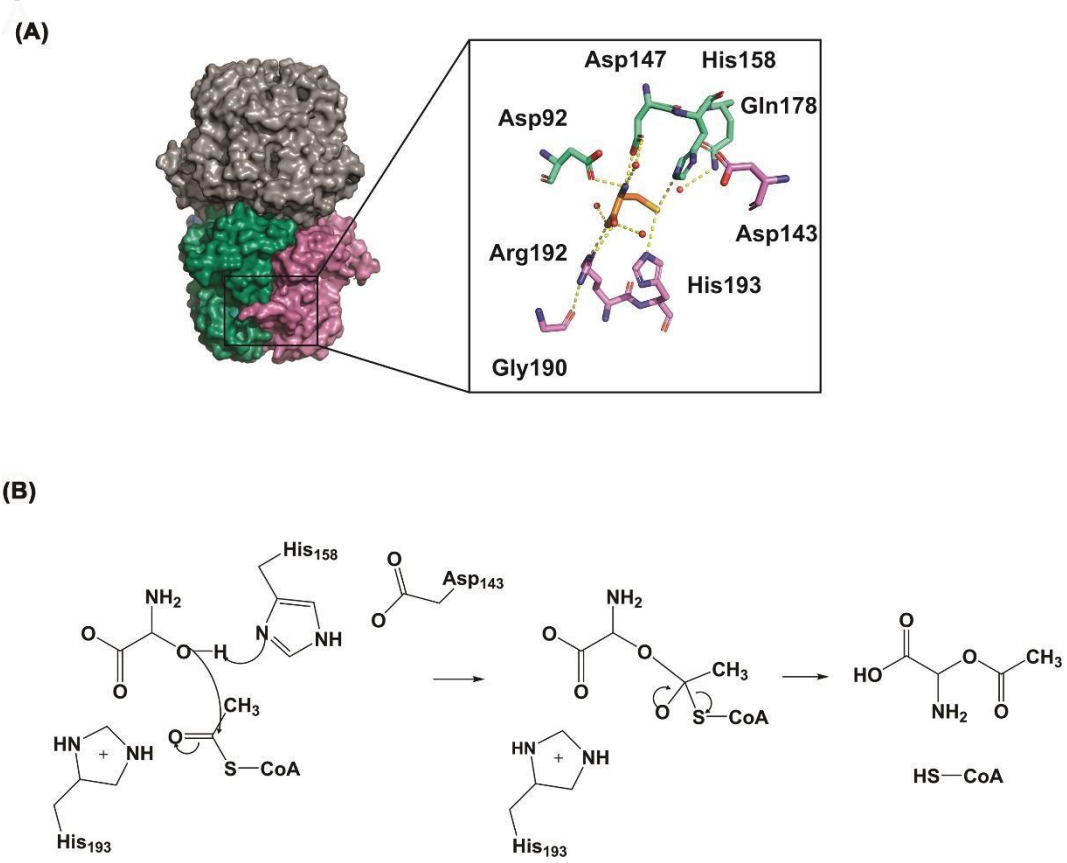


Figure 2.2. L-cysteine interactions with L-serine binding pocket in CysE from *E. coli* (1T3D). (A) Active site residues are represented as sticks, coloured green (Asp92, Asp147, His158 and Gln178) and pink (Asp143, Gly190, Arg192 and His193), based on chains. Inhibitor L-cysteine is represented as orange sticks. Hydrogen bonds are shown as yellow dashed lines. (B) CysE reaction mechanism for formation of *O*-acetylserine, adapted from (Pye et al., 2004). Figure was produced using PyMOL and ChemDraw Prime (RRID:SCR 016768).

2.5.4 Development of CysE Inhibitors

CysE enzymes from bacterial pathogens have been extensively characterised, both kinetically and structurally, with numerous high-quality crystal structurers available for inhibitor design. Given the importance of CysE in infection and antibiotic resistance, and its essentiality in some bacterial pathogens, CysE represents an attractive drug target. The inhibition of CysE would deplete the cell of cysteine and *O*-acetylserine, where the later isomerises to *N*-acetylserine, the natural inducer of the cysteine biosynthetic operon, leading to metabolic dysregulation. There have been limited studies into inhibitors of CysE enzymes, but promising inhibitors ($IC_{50} \leq 100$ μM) have been identified for a number of bacterial pathogens (Table 2.1) and are discussed below.

Table 2.1 List of characterised CysE inhibitors.

Inhibitor	**Enzyme**	**IC_{50} (μM)***	**K_i (μM) AcCoA***	**K_i (μM) L-serine**	**Citation**
Compound 3	*Ec*CysE	72[a]	42[a]	ND	Agarwal et al. (2008)
Compound 4	*Sa*CysE	71.84 ± 0.27	225.3[a]	53.9[a]	Chen et al. (2019)
Compound 30		71.84 ± 0.15	111.5[a]	47.66[a]	
Quercetin	*Kp*CysE	3.7[a]	162	ND	Verma et al. (2020)
Compound 3a	*St*CysE	48.6 ± 8.43	ND	ND	Magalhães et al. (2020)
Compound 5		110 ± 0	64 ± 12	ND	Magalhães et al. (2021)
Compound 22d		4.24 ± 0.11	ND	ND	Magalhães et al. (2021)

* Error reported as standard error

[a]No error reported

ND- Not determined

2.5.5 Natural Compound Inhibitors

Several promising CysE inhibitors have been identified via *in silico* screening of natural compound libraries. Recently, the flavonoid quercetin was found to inhibit CysE from *K. pneumoniae* (IC_{50} = 3.7 μM) (Table 2.1) (Verma et al., 2020). Through docking analysis, quercetin (Figure 2.3) was shown to bind allosterically to the CysE trimer–trimer interface. Although not experimentally investigated by the authors, the binding of quercetin to this interface may inhibit *Kp*CysE through disrupting the trimer–trimer interactions, dissociating the hexamer, which has been shown to reduce CysE activity (Verma et al., 2020). However, quercetin has been shown to inhibit other bacterial enzymes including isocitrate lyase (Shukla et al., 2015) and glutamate racemase (Pawar, Jha, Chopra, Chaudhry, & Saluja, 2020). This broad inhibition suggests that quercetin inhibition of CysE might be nonspecific, which is supported by the targeting of the trimeric interface and not serine or acetyl-CoA binding sites.

Natural compound inhibitors have also been identified for *Sa*CysE from methicillin-resistant *S. aureus* (MRSA) (Chen, Yan, et al., 2019). These include two polycyclic diterpenoids; compound 4 (11-oxo-ebracteolatanolide B) and compound 30 ((4R,4aR)-dihydroxy-3-hydroxymethyl-7,7,10a-trimethyl-2,4,4a,5,6,6a,7,8,9,10,10a,l0b-dodecahydrophenanthro[3,2-b]furan-2-one) (Figure 2.3). These compounds share the same chemical scaffold, with substitution of oxygens attached to the phenyl rings. Both compounds inhibited *Sa*CysE (both IC_{50} = 71.84 μM), where compound 4 was shown to display mixed inhibition against serine and competitive inhibition against acetyl-CoA, while the opposite was seen for compound 30. Docking analysis with a structural homology model of *Sa*CysE shows hydrophilic interactions between compound 4 and key catalytic residue, His95, and the N-terminal domain residues, Ala43 and Gly44. Compound 30 interacts with these identical residues and active site residue Asp94. Given the highly similar chemical structure of these compounds, it is unsurprising that they share CysE residue interactions; however what is interesting is these compounds do not dock in either the acetyl-CoA or serine binding pocket, but instead bind in a pocket between the third α-helix (equivalent to sixth α-helix in *E. coli*, 1T3D) and the serine binding site. This is a unique method of targeting the CysE active site as the inhibitor can interact with the active site residues without occupying the active

site, and may explain the mixed and competitive inhibition observed for these compounds.

Both compounds were also able to inhibit MRSA growth with minimum inhibitory concentration (MIC) values of 12.5 and 25 μg/ml for compound 4 and 30, respectively (Chen, Yan, et al., 2019). Furthermore, both compounds were able to disrupt a mature MRSA-biofilm at one-fold the MIC concentration and did not display any cytotoxicity to human cells. Both compounds did not inhibit the structurally similar hexapeptide enzyme, GlmU (N-acetylglucosamine-1-phosphate uridyltransferase) demonstrating target specificity. Although the *in vitro* IC_{50} values are relatively high, given that these compounds inhibited growth of the target organism, displayed target specificity and are well tolerated by mammalian cells, these compounds are ideal for further optimisation. Natural compounds have become popular for identifying novel antimicrobials, a large number of chemically distinct compounds can be screened and optimised by substituting chemical groups. Promising natural compound inhibitors discussed here could also be used for targeting other CysE homologues.

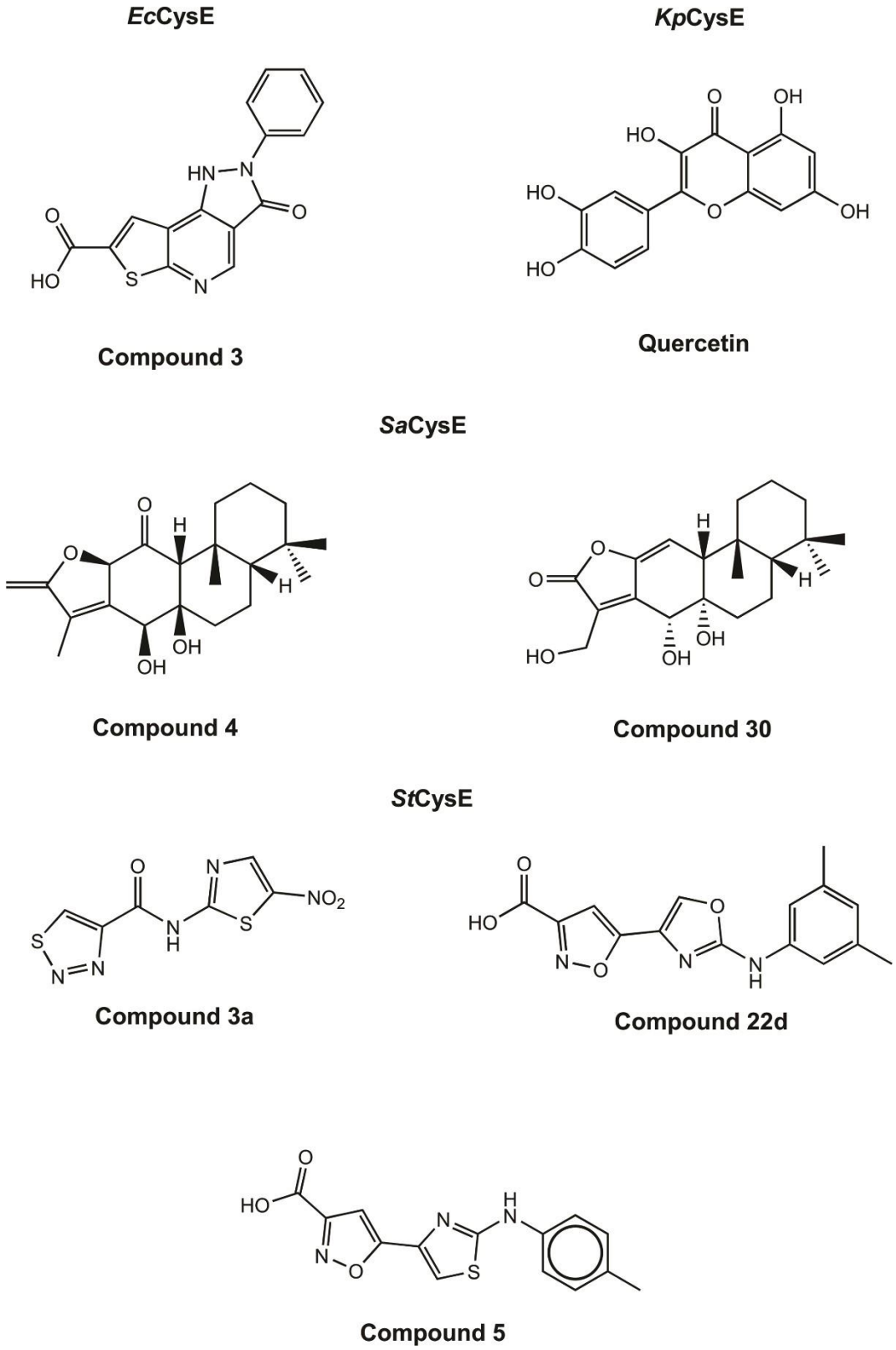


Figure 2.3 Chemical structures of CysE inhibitors. Figure produced using ChemDraw Prime (RRID:SCR 016768).

2.5.6 Chemical Inhibitors

Initial research into the development of CysE inhibitors was conducted by Agarwal et al. (2008) (Agarwal, Jain, Bhattacharya, & Azam, 2008). The researchers employed virtual screening to identify inhibitors of CysE from *E. coli*. Screening of the crystal structure of *Ec*CysE (1T3D), identified nine compounds with promising docking scores, of which three were characterised *in vitro*. Compound 3, (3-oxo-2-phenyl-3,5-dihydro-2H-pyrazolo(3,4d)thieno(2,3-b)pyridine-7carboxylic acid), was the only compound identified to inhibit *Ec*CysE (72 μM) and exhibit antimicrobial effects. This compound was tested for growth inhibition of the parasite *E. histolytica* resulting in unexpectedly potent inhibition (IC_{50} = 0.61 μM), suggesting off-target inhibition.

As well as natural products, 2-aminothiazole and 2-aminooxazole compounds have been investigated as inhibitors of CysE enzymes. These compounds mimic binding interactions with key active site residues similar to the natural inhibitor L-cysteine (Figure 2.2). Recent studies have explored 2-aminothiazoles and 2-aminooxazole compounds as inhibitors of CysE from *S. typhimurium* (*St*CysE) (Magalhães et al., 2021; Magalhães et al., 2020). Since the crystal structure for *St*CysE (7E3Y) has been solved only recently (Momitani, Shiba, Sawa, Ono, & Hurukawa, 2022), all virtual screening was carried out against both *Ec*CysE (1T3D) and *H. influenzae* CysE (1SSM) crystal structures, as there is strong conservation of active site residues with *St*CysE. Virtual screening of ~91,000 compounds from three libraries identified six compounds, which *in vitro* had IC_{50} values ranging from 13.6 to 84.1 μM (Magalhães et al., 2020). Further characterisation of these compounds, revealed only compound 3a (N-(5-Nitro-1,3-thiazol-2-yl)-1,2,3-thiadiazole-4-carboxamide), a 2-aminothiazole (Figure 2.3), to inhibit *St*CysE with an IC_{50} of 48.6 μM to be bactericidal, with an MIC of 64 μg/mL against *E. coli*. *In silico* docking analysis showed that compound 3a interacts with key active site residues Asp92, Asp157, Arg192, His193 and catalytic His158, mimicking interactions exhibited by the inhibitor L-cysteine (Figure 2.2). Compound 3a was shown to inhibit *E. coli* growth, but only in media low in cysteine. Previous research supports the anti-bacterial activity of 2-aminothiazoles, with bactericidal activity against *M. tuberculosis* reported (Kesicki et al., 2016; Meissner et al., 2013).

Further research into *St*CysE inhibition was conducted using an in-house library for further virtual screening of *St*CysE (Magalhães et al., 2021; Momitani et al., 2022). Using the same screening method as discussed previously the researchers identified seven compounds that reduced *St*CysE activity, with the most potent being the substituted 2-aminothiazole, compound 5, (5-{2-[(4-Methylphenyl)amino]-1,3thiazol-4-yl}-1,2-oxazole-3-carboxylic acid) (Figure 2.3) with an IC_{50} of 110 µM, which displayed competitive inhibition relative to acetyl-CoA (K_i = 64 µM). Docking analysis showed compound 5 interacts with the CysE serine active site residues and acetyl-CoA binding pocket, where the carboxylic acid group interacts with the same residues as seen for inhibitor cysteine/substrate serine. The 'L-shape' of the inhibitor allows it to mimic acetyl-CoA, explaining the observation of competitive inhibition (Magalhães et al., 2021). Structure–activity relationship analysis was undertaken through *in vitro* screening of compound 5 analogues. Substitution of the 2-aminothiazole ring with a 2-aminooxazole was shown to increase affinity, and the presence of an ester, amide or carboxylic acid group connected to the isoxazole ring was shown to be essential for affinity (Magalhães et al., 2021). Isoxazole-3-ester and isoxazole-3-carboxylic acid derivatives were further optimised through synthesis with different chemical groups connected to the oxazole ring (Magalhães et al., 2021). Affinity was not substantially affected by the side group, but derivatives with electron-withdrawing groups were unstable compared with electron-donating groups. The most potent analogue was compound 22d, (3,5-dimethylphenyl-(2-aminooxazol-4-yl) isoxazole-3- carboxylic Acid) (Figure 2.3), with an IC_{50} of 4.2 µM (Table 2.1). Unfortunately, this compound was unable to inhibit the growth of *E. coli*, requiring further optimisation to improve compound permeability.

Recently, Toyomoto and colleagues identified that alkyl gallates are potent inhibitors of *St*CysE (Toyomoto et al., 2023). Octyl gallate (OGA), exhibited *in vitro* inhibition of *St*CysE in the low micromolar range (IC_{50} = 3.6 µM) and showed antibacterial activity (IC_{50} = 59 µM) against *E. coli*. OGA treatment depleted intracellular thiol metabolite concentrations, including cysteine, and both reduced and oxidised forms of glutathione, comparable to a *cysE* deletion strain, as well as increasing susceptibility to oxidative stress. Furthermore, OGA treatment increased antibiotic sensitivity for antibiotic-resistant clinical isolates of *E. coli* and

K. pneumoniae (confirmed to both carry the metallo-β-lactamase gene, IMP-1) with reduced MICs reported for several carbapenem antibiotics, notably with a four-fold MIC reduction (imipenem and doripenem) seen for the *K. pneumoniae* isolate. These findings demonstrate that targeted inhibition of CysE could be an avenue for developing carbapenem adjuvants and could be used in the treatment of infections caused by carbapenem-resistant Gram-negative pathogens.

Overall, CysE inhibitor development is in its early stages, with a number of different strategies employed, with the main challenge in obtaining potent inhibitors that also inhibit bacterial growth. The essentiality of CysE in the notoriously antibiotic-resistant pathogens *S. aureus* and *N. gonorrhoeae* highlights CysE as an ideal target for antimicrobial development. With more research being undertaken in targeting cysteine biosynthetic enzymes, overcoming the challenge of finding compounds that are potent and yet specific, while being permeable to target organisms, will lead to the development of promising inhibitors.

2.6 *O*-acetylserine sulphydrylase (CysK/CysM)

O-acetylserine sulphydrylase is a pyridoxal 5′ phosphate (PLP) dependent enzyme that catalyses the second step of the L-cysteine biosynthesis reaction, combining *O*-acetylserine and a sulphur donor into cysteine. OASS is present in bacteria as two isoforms, OASS-A (CysK) that utilises sulphide for the synthesis of cysteine, and OASS-B (CysM) that utilises sulphide and thiosulphate. Bacteria with *cysK* or *cysM* deleted from the genome exhibit reduced virulence, compromised fitness and decreased antibiotic resistance. Subsequently, its inactivation is being pursued as a strategy for the identification of novel antibiotics and/or antibiotic adjuvants that target non-essential proteins.

2.6.1 Role and essentiality of CysK/CysM in bacterial pathogens

As the second and final enzyme in the cysteine biosynthetic pathway both OASS isoforms play an important role in bacteria. Transposon mutagenesis screening found that CysK is essential in just two bacterial pathogens, *Campylobacter jejuni* (Mandal, Jiang, & Kwon, 2017), and *Francisella novicida* (Gallagher et al., 2007). Like CysK, CysM is essential in very few bacterial species, including *Burkholderia pseudomallei* (Moule et al., 2014), and two strains of

Burkholderia cenocepacia, K56-2 (Gislason, Turner, Domaratzki, & Cardona, 2017), and J2315 (Wong et al., 2016). Given that many bacterial species have both OASS isoforms, or even two copies of CysK, it is not surprising CysK and/or CysM are non-essential in many bacterial pathogens. For example, the *M. tuberculosis* genome contains three annotated OASS genes, denoted CysK1, CysK2 and CysM. The nomenclature of these genes is confusing in that OASS-A is denoted as CysK1, OASS-B as CysK2 (not CysM) and the mycobacterial CysM is unique and found only in actinobacteria. *Mtb*CysM uses a small thiocarboxylated protein (CysO) as the sulphur donor and *O*-phosphoserine (not *O*-acetylserine) as the preferred acceptor substrate (Agren, Schnell, & Schneider, 2009; Burns-Huang & Mundhra, 2019; Burns et al., 2005; Claus, Zocher, Maier, & Schulz, 2005). Disrupting the *de novo* cysteine biosynthesis pathway in *M. tuberculosis* represents an attractive drug target. Cysteine biosynthesis is consistently upregulated in dormancy models of infection (Schnappinger et al., 2003; Voskuil, Visconti, & Schoolnik, 2004), particularly survival of *M. tuberculosis* in infected macrophages in the granuloma, where it is exposed to an extremely hostile environment. It could be argued that *M. tuberculosis* could obtain cysteine from the host and not be dependent on *de novo* synthesis of cysteine. However, the up-regulation of sulphur acquisition and cysteine synthesis genes in persister cells suggests that the host does not provide a sufficient amount of cysteine (Betts, Lukey, Robb, McAdam, & Duncan, 2002; Hampshire et al., 2004; Schnappinger et al., 2003; Voskuil, Bartek, Visconti, & Schoolnik, 2011; Voskuil et al., 2004), and it is likely cysteine is scarce within the granuloma due to host defense strategies such as nutrient depletion. *M. tuberculosis cysO* and *cysM* deletion strains show attenuation in *in vitro* models of dormancy and also in a mouse model of infection (Sassetti & Rubin, 2003). A target identification pipeline for drug targets in *M. tuberculosis* based on a comprehensive *in silico* analysis using experimental derived phenotype data and proteomics, suggests that both CysE and CysK2 are high confidence drug targets (Raman, Yeturu, & Chandra, 2008).

As discussed earlier, inactivation of the cysteine biosynthetic operon leading to cysteine auxotrophy in *S. typhimurium* led to an increased susceptibility to antibiotics during swarming, which is normally associated with a decreased susceptibility to antibiotics. A *S. typhimurium* Δ *cysk, cysM* double deletion strain

was pleiotropic (Turnbull & Surette, 2010), making it difficult to associate a particular phenotype to this strain, presumably due to the accumulation of toxic intermediates, such as 3′ phosphoadenoside 5′ -phosphosulphate (Neuwald et al., 1992).

Shatalin et al. (2011) linked both the high concentration of H_2S, and decreased cysteine concentration to increased resistance to a variety of antibiotics. H_2S is cytoprotective in some bacteria due to its ability to suppress oxidative stress generated by some antibiotics (Shatalin, Shatalina, Mironov, & Nudler, 2011). An *S. typhimurium cysK* deletion strain had decreased cysteine production, resulting in H_2S accumulation causing an eight-fold higher resistance to ofloxacin compared to wild-type (Frávega et al., 2016). This highlights the various roles of CysK in antibiotic susceptibility, enhancing resistance to some and decreasing resistance to others, due to the target of the antibiotic and degree of oxidative stress within the cell.

Metal ions at low concentrations are beneficial to bacteria, however, can become toxic at higher concentrations causing oxidative stress and eventually cell death (Joshi, Gupta, & Gupta, 2019). Studies in *S. typhimurium* LT2 and *E. coli* demonstrated that CysK plays an essential role in mediating resistance to the metal ion, tellurite (K_2TeO_3) (Ramírez, Castañeda, Xiqui, Sosa, & Baca, 2006; Vásquez, Saavedra, Loyola, Araya, & Pichuantes, 2001), which exhibits strong oxidising properties through an unknown mechanism. Deletion of *cysK* from *Azospirllum brasilense* conferred a lower MIC when grown on media with tellurite, whilst transfer of the *A. brasilense* and *Bacillus stearothermophilus cysK* into *E. coli* and *S. typhimurium* respectively, conferred increased tellurite resistance (Ramírez et al., 2006; Vásquez et al., 2001).

As mentioned previously, disruption of the cysteine biosynthetic pathways can affect biofilm formation. The effect of a number of mutants from the cysteine biosynthetic operon of *Vibrio fischeri*, including *ΔcysH*, *ΔcysJ*, *ΔcysK* and *ΔcysN* on biofilm formation were tested, with the greatest effect on biofilm formation seen with the *ΔcysK* mutant (Singh, Brooks, Ray, Mandel, & Visick, 2015). Biofilm and pellicle formation are vital to colonisation which was observed in early colonisation of baby squid where Δ*cysK* resulted in decreased colonisation (Singh et al., 2015).

Addition of cysteine allowed rescue of the biofilm defect and partial rescue of the pellicle defect, indicating a key role of CysK in initial colonisation (Singh et al., 2015).

In a screen for genes important for the switch of *N. meningitidis* from commensal to pathogen, CysK was identified as being important in this switch. Saturated random transposon insertion libraries of *N. meningitidis* were engineered and assessed for fitness during normal growth and colonisation of epithelial and endothelial cells, and the CysK mutant was identified as being of particular importance for epithelial cell infection (Capel et al., 2016).

2.6.2 Structural characteristics of CysK and CysM

O-acetylserine sulphydrylase (OASS; EC 2.5.1.47) belongs to the tryptophan synthase β-superfamily, and the β-family of PLP dependent enzymes (Guédon & Martin-Verstraete, 2007; Mino & Ishikawa, 2003; Takumi & Nonaka, 2016). PLP is an essential cofactor, utilised in the active site of OASS enzymes. As briefly discussed above there are two OASS isoforms, which each use an alternate sulphur source; CysK utilises sulphide, whereas, CysM utilises both thiosulphate and sulphide, with *O*-acetylserine to form cysteine (Kredich, 2008). Both isoforms are present in most bacteria, enabling the utilisation of both thiosulphate and sulphide for cysteine biosynthesis. Expression of these isoforms is influenced by aerobic or anaerobic growth, for example, the genome of the pathogen *S. typhimurium* encodes both OASS isoforms with CysK expressed in excess of CysM under aerobic conditions, and vice versa under anaerobic conditions (Filutowicz, Wiater, & Hulanicka, 1982). The two isozymes, CysM and CysK, function as homodimers and exhibit 25–45% similarity in their amino acid sequence (Agren et al., 2009; Schnell, Oehlmann, Singh, & Schneider, 2007; Tai, Nalabolu, Jacobson, Minter, & Cook, 1993; Zhao, Kumada, Imanaka, Imamura, & Nakanishi, 2006). The key difference between these isoenzymes is the ability of CysM to utilise larger sulphur donor substrates including thiosulphate, where CysK is only capable of accepting the small sulphur donor, hydrogen sulphide (Campanini, Benoni, et al., 2015; Claus et al., 2005; Tai et al., 1993).

CysK and CysM enzymes have been structurally characterised from a range of a bacterial species, enabling comparison of the two isoforms, and providing insight

into active site conformation and isoform differences for isoform specific inhibitor design. Within the CysK and CysM homodimer one PLP molecule is bound per subunit in the active site cleft that is formed between the N and C-terminal domains of each monomer. The active site cleft of CysK and CysM are fairly similar and are lined by seven chain segments totalling 16 residues. Six of these chain segments are highly conserved between CysK and CysM isoforms, however, the seventh segment spanning residues 210-216 of CysM (*E. coli* CysM numbering) indicates a key difference between the two enzymes (Chattopadhyay et al., 2007; Claus et al., 2005). CysM contains the three large residues Arg210, Arg211 and Trp212 followed by a three-residue insertion, which bulges toward the surface enlarging the active site. Most of this enlarged active site is occupied by the sidechain tail of Arg210 which most likely binds to thiosulphate or other larger sulphur donor substrates (Claus et al., 2005). Conversely, *St*CysK has three small residues, Gly230, Ala231 and Gly232 (Claus et al., 2005; Rabeh & Cook, 2004). This small change reduces the size of the active site cleft, therefore restricting CysK to using bisulphide as its sulphur donor. Lys41, is highly conserved across both CysK and CysM isozymes as it forms an internal aldimine linkage with PLP in both isoforms (Liang, Han, Tan, Ding, & Li, 2019; Rabeh & Cook, 2004). The enzyme cycles through open and closed conformations during catalysis. Both open (no substrates bound) and closed forms (substrates bound) have been structurally characterised (Burkhard et al., 1998; Burkhard, Tai, Ristroph, Cook, & Jansonius, 1999; Chattopadhyay et al., 2007) with the closed form occurring via binding of the substrate α-carboxylate or acetate to the substrate binding loop (residues 68–71 in *S. typhimurium*) triggering the active site to close (Burkhard et al., 1999; Chattopadhyay et al., 2007).

The two monomer subunits interact solely at the dimer interface and an allosteric binding site was recently identified in the CysK isoform, that is not present in CysM (Rabeh & Cook, 2004). The structure of CysK with Cl^- bound at the allosteric anion binding site at the subunit interface, shows a closed/inhibited form of the enzyme (Tai et al., 2001). The anion sulphide behaves similarly to chloride and has essentially the same ionic radii and also binds to the anion binding site (Tai et al., 2001), presumably acting as a further regulator of CysK activity and modulating sulphur flux within the cell.

2.6.3 CysK/CysM enzyme mechanism

As discussed for CysE, a detailed understanding of both structure and enzyme mechanism is important for inhibitor design. OASS catalyses the replacement of the β-acetoxy group of *O*-acetylserine by either sulphide in the case of CysK or thiosulphate and sulphide in the case of CysM to produce L-cysteine. Both CysM and CysK utilise a bi-bi ping-pong mechanism for cysteine biosynthesis (Tai et al., 1993; Zhao et al., 2006). This involves the release of one of the half-reactions products before all substrates have bound to the enzyme, generating an enzyme intermediate in the process (Ulusu, 2015). *O*-acetylserine carries out nucleophilic attack on the C-4′ of the internal aldimine (Figure 2.4). As the external aldimine is formed, the active site closes due to interaction of the substrate-carboxylate with an asparagine loop (Figure 2.4). A conserved lysine (Lys41, *E. coli* numbering), initially part of the internal aldimine PLP linkage, serves as a general base to deprotonate Cα in the β-elimination reaction that releases acetate at the conclusion of the first half-reaction forming the α-aminoacrylate intermediate (Figure 2.4) (Joshi et al., 2019; Rabeh & Cook, 2004). Acetate diffuses out of the active site as it opens partially to allow entry of bisulphide and product release. Lys41 remains protonated at the beginning of the second half-reaction, bisulphide diffuses into the active site attacking the Cβ of the α-aminoacrylate giving the cysteine external Schiff base (Figure 2.4). The cysteine product, *S*-sulphocysteine for CysM, and L-cysteine for CysK, is expelled via transimination (Joshi et al., 2019; Rabeh & Cook, 2004). *S*-sulphocysteine is reduced to L-cysteine via an unknown mechanism (Kredich, 2008).

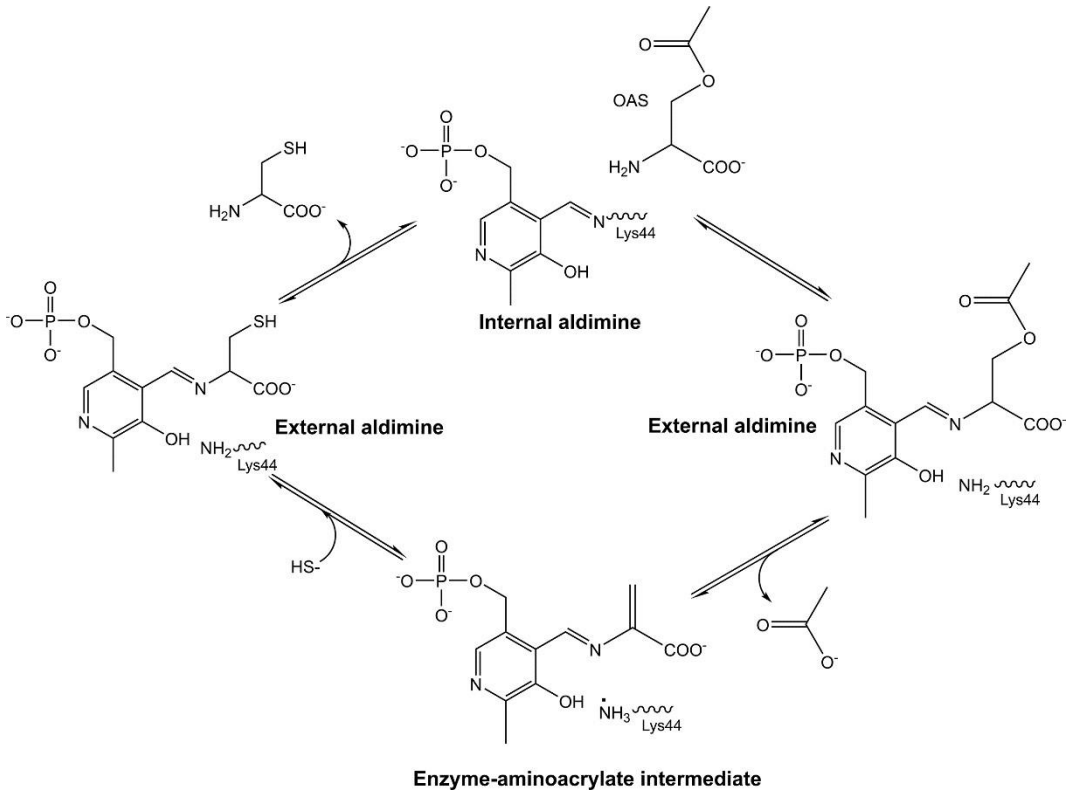


Figure 2.4 CysK catalytic mechanism, based on the *M. tuberculosis* CysK enzyme (*Mtb*CysK Lys44 equivalent to Lys41 in text). Figure produced using ChemDraw Prime (RRID:SCR_016768), adapted from (Joshi et al., 2019).

2.7 Cysteine synthase complex formation

The bienzyme cysteine synthase complex (CSC) was first discovered during the purification of S. typhimurium CysK and CysE (Kredich, Becker, & Tomkins, 1969; Kredich & Tomkins, 1966). CSC formation has since been confirmed across many other species including *E. coli* (Benoni, De Bei, et al., 2017; Mino, Hiraoka, et al., 2000; Mino, Yamanoue, et al., 2000), *H. influenzae* (Huang, Vetting, & Roderick, 2005; Salsi et al., 2010) and various plant species (Francois, Kumaran, & Jez, 2006; Yi et al., 2013). The formation of the CSC represents a unique avenue for the design of inhibitors. The complex forms via binding of the CysE C-terminal tail into the CysK active site, thus inhibiting CysK activity and therefore cysteine production (Benoni, De Bei, et al., 2017; Campanini, Pieroni, et al., 2015; Francois et al., 2006; Huang et al., 2005; Mino, Hiraoka, et al., 2000; Mino, Yamanoue, et al., 2000; Salsi et al., 2010). The complex forms in 3:2 CysE:CysK ratio consisting of one CysE hexamer and two CysK dimers (Huang et al., 2005). Deletion of 20 C-terminal residues from CysE results in an inability to bind to CysK and formation of the CSC does not occur, highlighting the importance of the C-terminal peptide tail of CysE

in CSC formation (Mino, Hiraoka, et al., 2000). CSC formation reduces L-cysteine feedback inhibition and L-serine substrate inhibition of CysE activity in *E. coli* (Benoni, De Bei, et al., 2017). Complex formation also reduces CysE cold inactivation at both 0 and 10°C (Mino, Yamanoue, et al., 2000), presumably due to increased stability caused by structural reorganisation at part of the N-terminal domain of CysE, that interacts with CysK leading to allosteric stabilisation at the interface between the CysE trimers (Rosa et al., 2021). CysM and CysE have no interactions and do not form a complex, due to differences in the active site structure (Zhao et al., 2006).

Structural studies of CysK in complex with CysE C-terminal peptides have provided insight into the interaction of the C-terminal CysE tail with the active site of CysK as to date there are no atomic structures of the CSC. The PLP cofactor in the OASS active site has fluorescent properties sensitive to its microenvironment and protein conformational changes. These fluorescent properties can be used to monitor formation of the CSC with CysE and with peptides that mimic the C-terminal tail of CysE (Campanini et al., 2005). Fluorescent monitoring of the PLP cofactor binding to the entire CysE protein and a C-terminal decapeptide (mimicking the CysE C-terminal tail), demonstrated that the C-terminal α-carboxylate of the CysE C-terminal decapeptide and the CysE C-terminal tail fit into the same position (Burkhard et al., 1999; Campanini et al., 2005).

At a ratio of 5:1 CysE:CysK (at which full complex formation is assumed to have taken place) the activity of CysK in the *E. coli* CSC is reduced to 10% of free CysK activity (Mino, Hiraoka, et al., 2000; Mino, Yamanoue, et al., 2000; Rosa et al., 2019). Yet the CysE C-terminal decapeptide when bound to CysK reduced activity to 50% at a 500:1 molar ratio of decapeptide to CysK (Mino, Hiraoka, et al., 2000). This can be attributed to full-length CysE binding 250-fold tighter to CysK compared with the C-terminal decapeptide (Huang et al., 2005). Dissociation constants (K_D) of peptides in complex with CysK compared with full-length CysE in complex with CysK further show the stark contrast in binding affinity of peptides compared with the full-length CysE in the CSC (Table 2.2). This indicates there may be additional structural features of CysE that CysK recognises aside from the C-terminal decapeptide, which increases the binding affinity but are not sufficient

for directing complex formation without the presence of the C-terminal decapeptide (Huang et al., 2005).

Table 2.2 Dissociation constants (KD) of CysE decapeptides and full-length CysE in the CSC. *Hi*DK = *Hi*CysE C-terminal decapeptide (GIDDGMNLNI) with *Hi*CysK from (Campanini et al., 2005). *Hi*DStK = *Hi*CysE C-terminal decapeptide (GIDDGMNLNI) with *St*CysK from (Campanini et al., 2005). *Ec*CSC1 = Full length *Ec*CysE with *Ec*CysK from (Benoni, De Bei, et al., 2017). *Ec*CSC2 = Full length *Ec*CysE with *Ec*CysK from (Marchetti et al., 2021). *St*CSC = Full length *St*CysE with *St*CysK from (Marchetti et al., 2021).

	CysE Decapeptide		**Full length CysE**		
	***Hi*DK**	***Hi*DStK**	***Ec*CSC1**	***Ec*CSC2**	***St*CSC**
K_D (nM)	515 ± 29*	972 ± 62*	4.5/6¤	0.63¤	0.83¤

*Error reported as standard error.

¤No error reported.

The highly conserved C-terminal isoleucine of the CysE C-terminal peptide is an essential anchor point for correct positioning of the C-terminal tail; in *H. influenzae* accounting for 80% of the total interaction energy (Salsi et al., 2010). This energy contribution is derived from Thr69 and Thr73 hydrogen bonds to the Ile267 α-carboxylate (Figure 2.5) (Salsi et al., 2010). The Ile267 sidechain forms hydrophobic interactions in an apolar pocket formed by the Phe144 and PLP (Figure 2.5) (Salsi et al., 2010). A further 10% of the total interaction energy is split equally between the C-terminal Asn266 and Leu265 (Figure 2.5) (Salsi et al., 2010).

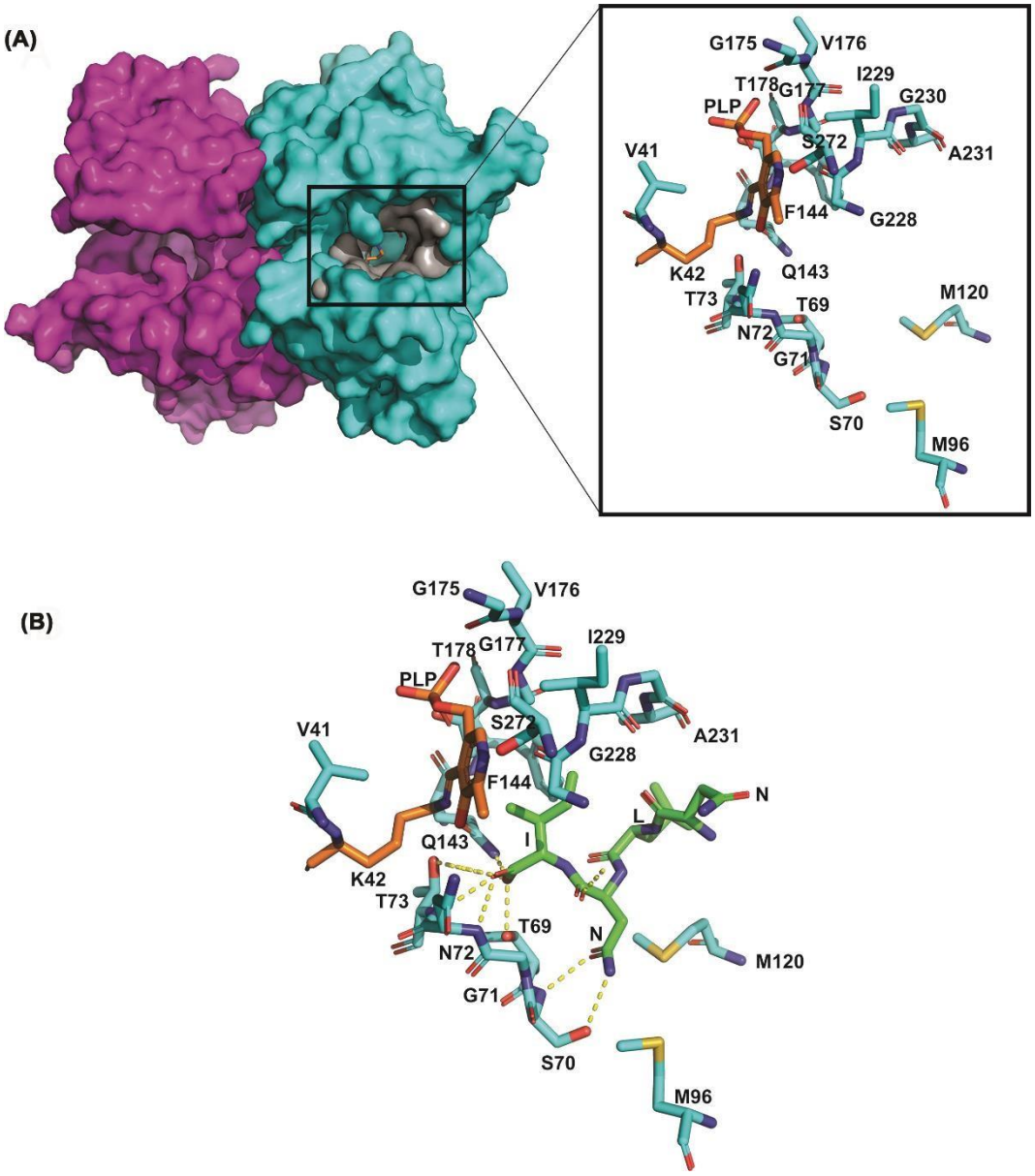


Figure 2.5 Overview of *H. influenzae* CysK. (A) *Hi*CysK dimer surface shown with one monomer in magenta and the other in cyan with PLP bound to K42 visible in orange deep in the active site cleft. Zoomed into *H. influenzae* active site residues (V41, K42, 69- TSGNT-73, M96, M120, 143-QF-144, 175-GVGT-178, 228GIGA-231 and S272) shown in cyan with PLP bound to K42 in orange. Active site is similar to *St*CysK open (0.5 Å, r.m.s.d) and *St*CysK partially closed anion inhibited (0.4 Å, r.m.s.d) conformation (Huang et al., 2005) (B) *H. influenzae* active site residues shown in cyan interacting with bound *Hi*CysE C-terminal tetrapeptide shown in green (NLNI). Polar bonds shown with dotted yellow lines. Figure created with 1Y7L from (Huang et al., 2005) using PyMOL.

The dependence of the CSC on sulphur availability indicates regulation of sulphur flux to be the purpose of CSC formation (Benoni, De Bei, et al., 2017; Kredich et al., 1969; Wang & Leyh, 2012; Zhao et al., 2006). The regulation forms a loop beginning with high availability of sulphur to the cell where the CSC is stabilised

by bisulphide (Kredich et al., 1969); however, when sulphur availability is low, OAS accumulates via CysE production, thus indicating sulphur starvation and dissociation of the CSC (Benoni, De Bei, et al., 2017; Kredich et al., 1969; Wang & Leyh, 2012; Chunhui Zhao et al., 2006). Dissociation of the CSC can occur at OAS concentrations upward of 50 μM (Kredich et al., 1969; Wang & Leyh, 2012), which then non-enzymatically isomerises to *N*-acetylserine and binds to the transcriptional regulator CysB, thus promoting expression of sulphate acquisition genes (Jovanovic, Lilic, Savic, & Jovanovic, 2003; Kredich et al., 1969; Wang & Leyh, 2012). Expression of sulphate acquisition and reduction genes increases the concentration of sulphur within the cell, completing the loop with high sulphur availability promoting CSC stability and therefore increasing OAS production.

2.8 Development of inhibitors for the OASS isoforms CysK and CysM

2.8.1 Peptide inhibitors

Salsi et al. (2010) and Spyrakis et al. (2013) have paved the way in CysK inhibitor discovery with the identification of multiple potent peptide inhibitors for several key CysK isoforms (Salsi et al., 2010; Spyrakis, Singh, et al., 2013). Although, more recently, CysK inhibitor studies have focused primarily on chemical inhibitors, these peptide inhibitors stand as key templates for chemical inhibitor designs. The natural inhibition of CysK by CysE has been routinely used for these peptide inhibitor studies as a design platform (Campanini, Pieroni, et al., 2015; Salsi et al., 2010; Spyrakis, Felici, et al., 2013; Spyrakis, Singh, et al., 2013). The three C-terminal residues of CysE contribute the strongest interactions with CysK, and therefore a minimum three-residue peptide is required for CysK inhibition; both Salsi et al. (2010) and Spyrakis et al. (2013) screened pentapeptides—including an additional two residues to better mimic the full-length CysE C-terminal peptide (Salsi et al., 2010; Spyrakis, Felici, et al., 2013; Spyrakis, Singh, et al., 2013). Salsi et al. (2010) utilised the *Hi*CysK crystal structure (1Y7L) in complex with the *Hi*CysE C-terminal peptide (Huang et al., 2005), replacing this peptide with a panel of 400 pentapeptides into the active site via virtual screening (Salsi et al., 2010). Spyrakis et al. (2013) followed through an analogous computational pipeline, with

the inclusion of the *S. typhimurium* apo CysK (*St*CysK) (1OAS) (Burkhard et al., 1999) and *St*CysM (2JC3) (Chattopadhyay et al., 2007) crystal structures.

Peptide inhibitors for both *S. typhimurium* isoforms and *Hi*CysK were identified, which demonstrated improved potencies *in vitro* compared with their respective native CysE C-terminal pentapeptides (Table 2.3). The most potent pentapeptides identified for *Hi*CysK were MNWNI and MNYDI, which both exhibited approximately 1.75 times improved affinity for the enzyme in comparison to the equivalent native *Hi*CysE pentapeptide (MNLNI). Interestingly, structural analysis (MNWNI, 3IQG; MNYDI, 3IQH) showed the asparagine in MNWNI at position four to participate in hydrogen-bonding with Ser70; whereas, the aspartate of MNYDI was shown to protrude out toward the protein surface. This protrusion is thought to be a result of the tyrosine at position three participating in an aromatic cluster with Phe144 and Phe233, which allows a sulphate ion to reside within the active site instead, and mimic the interactions of the asparagine in MNWNI (Figure 2.6) (Salsi et al., 2010).

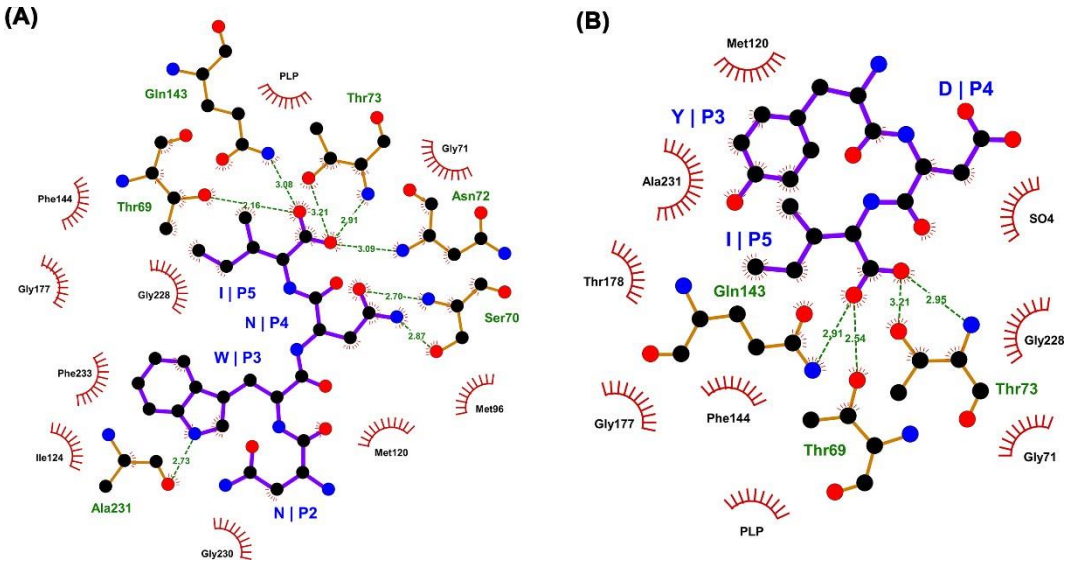


Figure 2.6 Interaction of MNWNI and MYDI peptides with *Hi*CysK. (A) LigPlot showing the interactions between the *Hi*CysK residues and the MNWNI pentapeptide (PDB code: 3IQG). (B) LigPlot showing the interactions between the *Hi*CysK residues and the MNYDI pentapeptide (PDB code: 3IQH).

Table 2.3 List of top characterised CysK peptide inhibitors.

Inhibitor	Enzyme	K_D (μM)*	Citation

MNLNI (*Hi*SAT WT)	*Hi*CysK	44.0 ± 3.6	Salsi et al. (2010)
	*St*CysK	120 ± 12	Spyrakis et al. (2013)
	*St*CysM	~3000¤	Spyrakis et al. (2013)
MNWNI	*Hi*CysK	24.9 ± 3.6	Salsi et al. (2010)
	*St*CysK	10.4 ± 0.9	Spyrakis et al. (2013)
MNYDI	*Hi*CysK	25.8 ± 1.7	Salsi et al. (2010)
	*St*CysK	0.22 ± 0.04	Spyrakis et al. (2013)
YGDGI (*St*SAT WT)	*St*CysK	11.8 ± 0.6	Spyrakis et al. (2013)
	*St*CysM	4922 ± 1,030	Spyrakis et al. (2013)
YGYDI	*St*CysK	0.42 ± 0.02	Spyrakis et al. (2013)
MNDGI	*St*CysK	306 ± 17	Spyrakis et al. (2013)
	*St*CysM	1100 ± 100	Spyrakis et al. (2013)

*Error reported as standard error
¤No error reported

Second, and most intriguingly, the top pentapeptide assessed for *St*CysK was MNYDI, which is a pentapeptide based on the *Hi*CysE C-terminus (MNLNI), and not that of the *St*CysE C-terminus (YGDGI) (Salsi et al., 2010). MNYDI showed approximately 60 times improved affinity compared with YGDGI, and approximately 600 times improved affinity compared with MNLNI. Moreover, the equivalent pentapeptide of MNYDI based on the *St*CysE sequence, YGYDI, still

showed reduced potency compared with MNYDI, highlighting that the terminal three residues of the pentapeptides are involved in affinity regulation, whereas the preceding two residues are involved in selectivity (Salsi et al., 2010).

Furthermore, the pentapeptide which demonstrated the greatest potency toward *St*CysM was based on both the *St* and *Hi*CysE C-terminal sequences—MNDGI. This pentapeptide exhibited approximately 4.5 times greater affinity for the enzyme than YGDGI, and around three times greater affinity than MNLNI (Salsi et al., 2010). This inhibitor is also the most effective against both *S. typhimurium* isoforms, with the difference in K_D values minimised. Correlation analysis of the K_D values of *St* isoform binders, demonstrated an inversely proportional relationship, where a more potent inhibitor of *St*CysK was more likely to possess reduced affinity for *St*CysM.

Reasonable affinity correlation was also noted between binders of CysK *H. influenzae* and *S. typhimurium* homologues. Comparison of the active site architecture of these homologues signified only one minor difference—the orientation of the Gln227 side chain. In *Hi*CysK, the R-group protrudes into the active site; whereas, in *St*CysK this group protrudes toward the enzyme surface. Consequently, it is proposed that this localises the third and fourth residues of the pentapeptides differentially within the enzyme active sites. Nevertheless, this highlights the possibility for synthesising broad-spectrum CysK and CysM compounds (Spyrakis, Singh, et al., 2013).

Altogether, these data demonstrate that effective peptide inhibitors of CysK should ideally possess negatively charged, hydrogen-bond acceptors at position four, and hydrophobic residues at position three. Unfortunately, this trend does not seem to translate to CysM peptide inhibitors, with glycine, a neutral and non-hydrogen bond acceptor at position four, and a negatively charged, hydrophilic residue at position three.

More recently, Kant et al. (2019) investigated a panel of tetrapeptides for inhibition of the parasite *Leishmania donovani* CysK (*Ld*CysK) (Kant et al., 2019), with the aim to deconvolute the findings of (Raj, Kumar, & Gourinath, 2012), where *Ld*CysK did not demonstrate a preference for tetrapeptides with either long or small residues. Tetrapeptides were designed to contain all possible amino acid

combinations for subsequent docking analysis. Docking analysis revealed EWSI and DWSI as the top two binders, respectively, with EWSI observing greater hydrogen-bonding and hence stabilisation capabilities (Kant et al., 2019). Therefore, EWSI stands as a starting point for future *Ld*CysK inhibitor designs, alongside the need for *in vitro* characterisation. Kant et al. (2019) also followed on to compare the differences between these identified tetrapeptide inhibitors of *Ld*CysK for their affinity for *Mtb*CysK1, given the similarity of these peptides to the native *Mtb*CysE tetrapeptide—DFSI. Interestingly, EWSI demonstrated improved docking into *Mtb*CysK1 compared with both DFSI and EFSI (Kant et al., 2019). This highlights EWSI as a valid starting point for future virtual screening of *Mtb*CysK1 chemical inhibitors.

2.8.2 Chemical inhibitors

Chemical inhibitors stand as superior drug compounds given their improved *in vivo* half-life, bioavailability, and pharmacokinetics compared with peptidic compounds (Amori et al., 2012; Reichelt & Martin, 2006). Since the studies by Salsi (2010) was published, many groups have been working on the design, synthesis and characterisation of chemical inhibitors for both CysK and CysM. The design work has largely stemmed from structurally mimicking the key chemical groups of the peptide inhibitors. Subsequently, many different chemical inhibitors have been identified, including, cyclopropane carboxylic acids (Amori et al., 2012; Annunziato et al., 2016; Annunziato et al., 2021; Bruno, Amori, & Costantino, 2013; Pieroni et al., 2016), thiazolidinone and pyrimidinone derivatives (Jean Kumar et al., 2013; Kumar, Raj, Nagpal, Subbarao, & Gourinath, 2011; Poyraz et al., 2013), fluoroalanines (Franko et al., 2018), benzoic acids (Brunner et al., 2016), and pyrazinamines and acetamides (Kant et al., 2019). The majority of these compounds were discovered via *in silico* docking experiments using crystal structures of inhibited complexes of the enzymes and subsequent high-throughput virtual screening.

To date, Amori et al. (2012) and Pieroni et al. (2016) have discovered the most potent chemical inhibitors of *Hi*CysK - (±)-trans-2-(ethoxycarbonyl) cyclopropanecarboxylic acid (±)-7 (referred to as UPAR40) (Amori et al., 2012) and trans-2-(prop-1-enyl)-cyclopropanecarboxylic acid (referred to as cyclopropyl

derivative 17) (Pieroni et al., 2016), respectively. Both compounds contain a carboxylate and a hydrophobic moiety which reflect the key properties of the CysE C-terminal isoleucine, and are incorporated together between a cyclopropane spacer (Figure 2.7) (cyclopropane spaces are common features in bioactive molecules and also aid in restraining the ligand in its effective conformation for enzymatic interaction) (Amori et al., 2012; Campanini, Pieroni, et al., 2015). Promisingly, docking and molecular dynamic analyses showed that these inhibitors lock the enzyme in its closed conformation (Amori et al., 2012; Bruno et al., 2013). Despite this success, Pieroni et al. (2016) has noted the impracticality of the chemical properties of these compounds for drug-like synthesis (Pieroni et al., 2016), and given the inactivity against the CysM isoform, efforts should now be directed toward improving the synthetic feasibility of these compounds, and their activity towards *Hi*CysM, alongside *in vivo* and cytotoxicity assays.

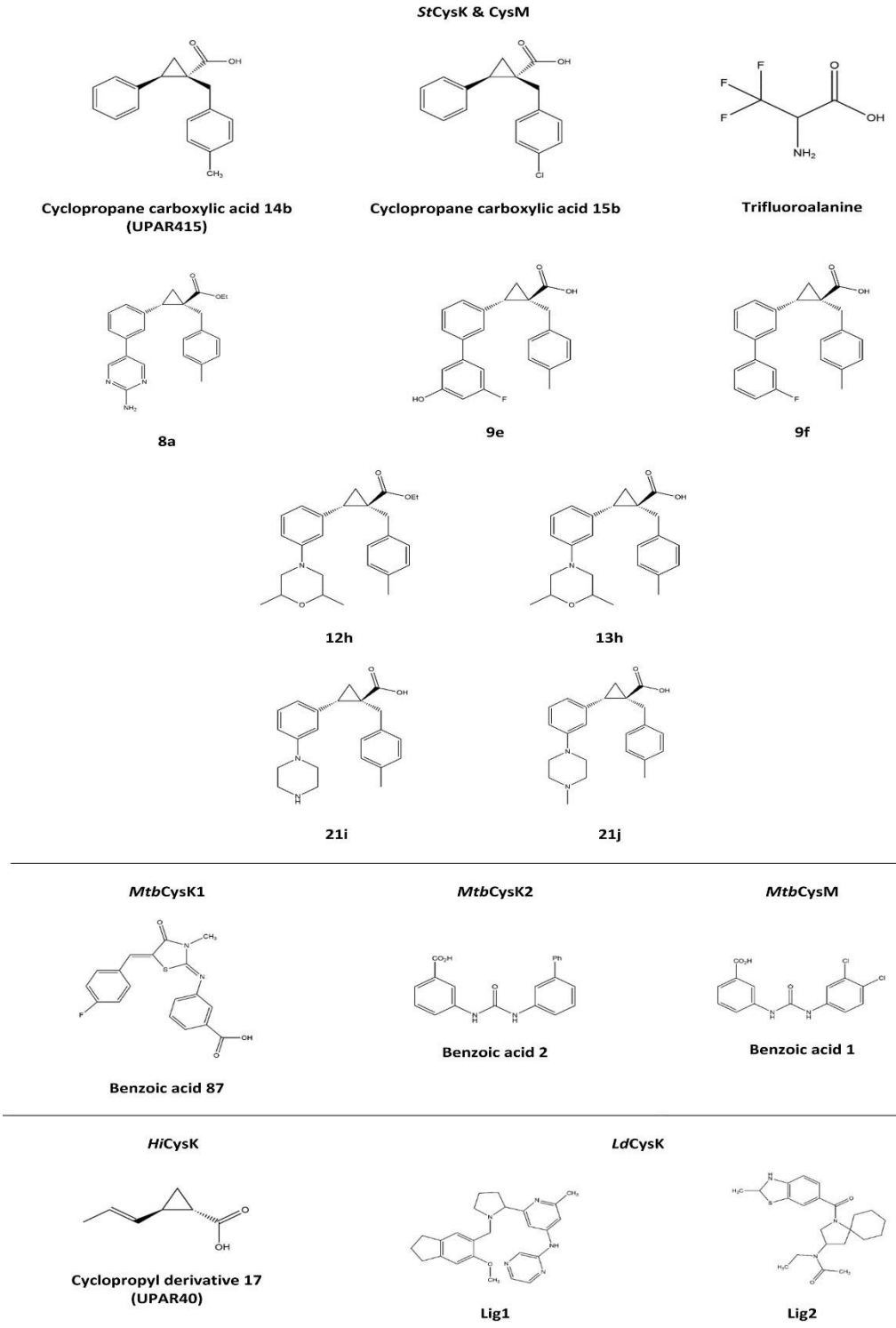


Figure 2.7 Chemical structures of top OASS chemical inhibitors. Figure produced using ChemDraw Prime (RRID:SCR_016768).

The success of the cyclopropane carboxylic acid compounds for *Hi*CysK, has also been shown for both *S. typhimurium* isoforms. Pieroni et al. (2016) identified (1S,2S)-1-(4-Methylbenzyl)-2-phenylcyclopropanecarboxylic acid 14b and (1S,2S)-1-(4-Chlorobenzyl)-2-phenylcyclopropanecarboxylic acid 15b as the most

potent inhibitors against *St*CysM to date (Table 2.4) These compounds were adapted from cyclopropyl derivative 17 to contain a phenyl group as opposed to a vinyl group (Figure 2.7), which better embodies the trans orientation with the carboxylate displayed by the CysE C-terminal isoleucine, and is more synthetically viable. It was noted that further substitution of cyclopropyl derivative 17 at the αcarbon position for interaction with both *S. typhimurium* isoforms would better occupy a moderately polar area of the active site, and therefore, benzyl substitution at this site with further para substitution resulted in compounds 14b and 15b (Pieroni et al., 2016). Interestingly, the para substitution of compounds 14b (4-CH_3) and 15b (4-Cl) represent both electron withdrawing and donating groups, yet reasonably equivalent potencies are observed (Table 2.4); however, compound 15b shows reduced selectivity toward either isoform. Promisingly, the IC_{50} values of compound 15b for both enzyme isoforms (Table 2.4), in the presence of the natural substrates of the enzyme, corroborate with the dissociation constants reported, indicating competitive inhibition.

Table 2.4 Top characterised OASS chemical inhibitors.

Inhibitor	**Enzyme**	**IC_{50} (µM)***	**K_D (µM)***	**Citation**
Cyclopropyl derivative 7 (UPAR40)	*Hi*CysK	700 ± 53	1.46 ± 0.25	Amori et al. (2012)
Cyclopropyl derivative 17			1.45[a]	Pieroni et al. (2016)
Cyclopropane carboxylic acid 14b (UPAR415)	*St*CysK		0.028 ± 0.005	
	*St*CysM		0.49 ± 0.05	
Cyclopropane carboxylic acid 15b	*St*CysK	0.099 ± 0.004	0.054 ± 0.008	
	*St*CysM	0.50 ± 0.03	0.42 ± 0.06	
8a	*St*CysK	ND	-	Annunziato et al. (2022)
	*St*CysM	ND	-	
9e	*St*CysK	ND	0.035 ± 0.003	
	*St*CysM	ND	0.61 ± 0.08	

9f	*St*CysK	ND	0.051 ± 0.004	
	*St*CysM	ND	1.45 ± 0.31	
12h	*St*CysK	ND	-	
	*St*CysM	ND	-	
13h	*St*CysK	ND	0.066 ± 0.005	
	*St*CysM	ND	3.37 ± 0.72	
21i	*St*CysK	ND	0.045 ± 0.09	
	*St*CysM	ND	83.8 ± 16.1	
21j	*St*CysK	ND	0.25 ± 0.06	
	*St*CysM	ND	23.6 ± 4.5	
Trifluoroalanine	*St*CysK	130 ± 10	ND	Franko et al. (2018)
	*St*CysM	1,290 ± 230	ND	
Benzoic acid 87	*Mtb*CysK1	0.019 ± 0.0011	ND	Poyraz et al. (2013)
Benzoic acid 1	*Mtb*CysM	ND	0.32 ± 0.01	Brunner et al. (2016)
Benzoic acid 2	*Mtb*OCysK2	ND	22.6 ± 2.4	

*Error reported as standard error.
[a] No error reported.

Recently, Annunziato et al. (2021) utilised compound 14b, here referred to as UPAR415, as an effective adjuvant for the polymyxin antibiotic, colistin (Annunziato et al., 2021). Interestingly, in the presence of low cysteine levels, administration of UPAR415 alone did not exhibit any bactericidal or bacteriostatic effects on multiple bacterial species (Gram-positive and -negative) (Annunziato et al., 2021). In contrast, when UPAR415 was treated under these same conditions in conjunction with colistin, significant deductions were seen in the MIC of colistin compared with when colistin is dosed on its own. Promisingly, the cytotoxicity of UPAR415 was also shown to be insignificant. The crystal structure of *St*CysK in complex with UPAR415 (6Z4N) was also solved by Annunziato et al. (2021). This demonstrated that UPAR415 is a competitive inhibitor of *St*CysK against its first substrate, *O*-acetylserine, and was found localised in proximity to the PLP cofactor

(Figure 2.8). The active site entrance could be seen to be partially blocked by the two aromatic substituents of the cyclopropane-ring, which engaged in hydrophobic interactions with active site residues. The remainder of the UPAR415 molecule was seen to penetrate into the active site. Promisingly, the carboxylate group of UPAR415 was found to localise to the well-studied carboxylic site; engagement with this site has been found to result in a conformational change of the enzyme into its closed active site state through translocation of the substrate-binding loop, which ultimately rotates the N-terminal domain over the active site. Although, UPAR415 positions itself at this carboxylic site, only partial closure of the active site is induced, which can be explained by a steric clash of the substrate-binding loop with the tolyl substituent of UPAR415 (Figure 2.8) (Annunziato et al., 2021).

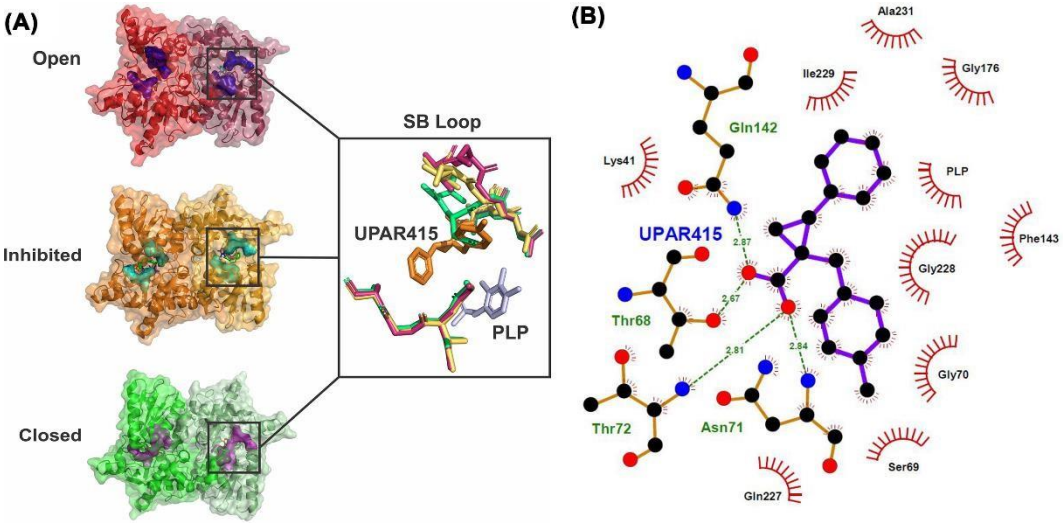


Figure 2.8 Inhibition of *St*CysK by UPAR415. (A) Structural comparison of the substrate binding loop between the open conformation (1OAS) shown in shades of red, the inhibited conformation (6Z4N) shown in shades of orange, and the closed conformation (1D6S) shown in shades of orange, and the closed conformation (1D6S) shown in shades of green. (B) LigPlot showing the interactions between the enzyme residues and the UPAR415 molecule. Figure produced using PyMOl and LigPlot.

A medicinal chemistry campaign to synthesise UPAR415 derivatives and the characterisation of ligand – target complexes revealed the presence of an accessory sub-pocket that can be filled by substitutions at the 3′ position of UPAR415 (Annunziato et al., 2022). Most compounds synthesised in this study retained good binding *in vitro*. Derivatives with heteroaliphatic or heteroaromatic groups in the 3′ position led improved inhibition against both OASS isoforms compared with

those previously described (Annunziato et al., 2021). Compounds substituted with a heteroaromatic group at the 3′ position for example, trans-2-(3′-fluoro-5′ hydroxy-[1,1′-biphenyl]-3-yl)-1-(4-methylbenzyl)cyclopropane carboxylic acid (compound 9e) and trans-2-(3′-fluoro-[1,1′-biphenyl]-3-yl)-1-(4methylbenzyl)cyclopropane carboxylic acid (compound 9f), that had a phenyl substituted with a fluorine and hydroxy group, or a fluorine respectively demonstrated potent inhibition (nanomolar range) against CysK (Figure 2.7) (Annunziato et al., 2022). Whereas compounds carrying a 2-aminopyrimidine group at the 3′ position showed good inhibition of both isoforms. Compound 13h (trans-2-(3-(2,6-dimethylmorpholino)phenyl)-1-(4methylbenzyl)cyclopropanecarboxylic acid) had a dimethyl morpholine at the 3′ position and had the most potent activity for derivatives substituted with a heteroaliphatic group at position 3′ (Figure 2.7). Compounds with a piperazine ring had good activity against CysK but these derivatives were less effective at inhibiting CysM. The majority of all compounds synthesised demonstrated low toxicity in that they were tolerated by mammalian cells. Compound 13h and its corresponding ester derivative 12 h showed potent inhibition *in vitro* and good toxicity profiles. These lead compounds were tested as colistin adjuvants, showing effective synergy by reducing the MIC of colistin against *E. coli* and *S. typhimurium* even at low concentrations (Annunziato et al., 2022). Importantly for this study, the most promising derivative 12h (trans-ethyl 2-(3-(2,6-dimethylmorpholino)phenyl)1-(4-methylbenzyl)cyclopropanecarboxylate) was linked to the chemical inhibition of CysK and CysM *in vivo* using target engagement experiments in *S. typhimurium* in the presence of colistin (the compound on its own does not exert any bactericidal effects) (Annunziato et al., 2022). Compound 12h demonstrated a significant improvement in the active concentration at which it can act as a colistin adjuvant inside cells, thereby paving the way as a prodrug to overcome some of the toxicity issues associated with colistin.

Franko et al. (2018) studied fluoroalanine derivatives *in vitro* against both *S. typhimurium* isoforms as irreversible inhibitors, given the ubiquitous use and study of these halogens as irreversible inhibitors of PLP-dependent enzymes

(Alexeev et al., 2006; Alston, Muramatsu, Ueda, & Bright, 1981; Azam & Jayaram, 2016; John & Tudball, 1972; Passera et al., 2011; Phillips & Dua, 1992; Silverman, 1995; Silverman & Abeles, 1976; Tysoe & Withers, 2014). Halogenated alanine addition to this group of enzymes is known to generate an unsaturated α-aminoacrylate Schiff's base (or α-aminoacrylate-PLP complex) (Alexeev et al., 2006; Poulin, Lu, Ackermann, Bey, & Pegg, 1992)—the subsequent inhibition mechanism is dependent on the specific halogenated alanine and how this reorientates the active site; typically, the catalytic lysine is attacked, and subsequent chemical rearrangement culminates in the disengagement of the halogen ions, allowing for a stable derivative to form, which ultimately inhibits the enzyme (Franko et al., 2018).

β,β,β-Trifluoroalanine (triF-Ala), which is a known PLP-dependent enzyme suicide inhibitor (Alexeev et al., 2006; Faraci & Walsh, 1989; Phillips & Dua, 1992; Silverman & Abeles, 1976, 1977; Tysoe & Withers, 2014; Wang & Walsh, 1981), was the most potent fluoroalanine derivative assessed. This compound had high affinity for the *St*CysK enzyme active site and slightly lower for that of *St*CysM. The mechanism of inhibition was determined to be irreversible covalent modification of the catalytic amino acids, yet the IC_{50} values of triF-Ala with both enzyme isoforms (Table 2.4) are too high to be an efficient inhibitor.

Poyraz et al. (2013) identified the most potent inhibitor to date for the *M. tuberculosis* enzyme CysK1 (*Mtb*CysK1), a thiazolidinone derivative -3-((Z)((Z)-5-(4-fluorobenzylidene)-3-methyl-4-oxothiazolidin-2-ylidene) amino) benzoic acid (Poyraz et al., 2013). E-pharmacophore sites were identified from the *Mtb*CysK1-DFSI crystal structure (2Q3C) (Schnell et al., 2007)—one aromatic ring, two acceptors, and two negative ionisable moieties. The compound benzoic acid 87 was identified from substitution analyses from the initial *in vitro* hit compound -3-({5-[2-(carboxymethoxy)benzylidene]-3-methyl-4-oxo-1,3thiazolidin-2-ylidene}amino)benzoic acid 2, where compound 87 represents a C4 fluoro substitution, and showed an approximately five fold greater inhibition effect compared to benzoic acid 2.

The crystal structure of derivative 2 bound to *Mtb*CysK1 (3ZEI) showed the thiazolidinone moiety to not protrude greatly out of the active site as is seen with the native peptide (Poyraz et al., 2013). The thiazolidine core mimics the phenyl group of the DFSI peptide and contributes to interactions with the enzyme hydrophobic cleft. Although similarly, the carboxylic moiety of the compound benzoic acid 2 (associated with the benzoic acid group) was found docked entirely within the active site with Ser72 interactions and potential for hydrogen-bonding with Thr71 and Gln144, as has previously been shown with the C-terminal isoleucines of peptide inhibitors (Salsi et al., 2010; Spyrakis, Felici, et al., 2013). In contrast with previous inhibitor observations, the 2-carboxymethoxy moiety does not participate in any solvent hydrogen-bonding despite protruding out toward the protein surface, which may indicate a key feature in developing future potent inhibitors (Campanini, Pieroni, et al., 2015; Poyraz et al., 2013). Nevertheless, this structural analysis revealed an enzyme pocket next to the para-position of the benzylidene ring, from which compound 87 verifies (Poyraz et al., 2013).

Brunner et al. (2016) identified the most potent inhibitor to date for *Mtb*CysM -3-(3-(3,4-dichlorophenyl)ureido) benzoic acid 1, as well as that for *Mtb*CysK2 - 3-(3([1,1-Biphenyl]-3-yl)ureido)benzoic acid 2 (Brunner et al., 2016). These hit compounds were identified by *in vitro* screening of approximately 17,000 small molecules, followed by structural analysis. The crystal structure of the compound benzoic acid 1 bound to *Mtb*CysM (5I7A) did not demonstrate large differences compared with the apo structure—the compound was shown to be bound within the open state of the active site (parallel to PLP pyrimidine ring plane), where stacking interactions were seen between the PLP pyrimidine ring and the urea moiety (Brunner et al., 2016). The meta carboxylate group of the compound was found associated within the active site, similar to the carboxylate moiety of the α-aminoacrylate intermediate of *Mtb*CysK1 (Schnell et al., 2007); whereas the 1,2dichlorobenzene group interacts within the hydrophobic cleft of the enzyme. The core urea group was found to interact with the Asn221 side chain via its carbonyl moiety, interactions with the carboxyl of Ala323 found within the active site occur through an amide group, with the other amide group forming a hydrogen-bond with a water molecule (Brunner et al., 2016). Promisingly, in a nutrient starvation model (simulating dormancy, when CysM is primarily expressed), compound 1

demonstrated higher potency compared with current clinically approved first-line tuberculosis antibiotics, with insignificant cytotoxic effects on several mammalian cell lines (Brunner et al., 2016).

Kant et al. (2019) identified two hit molecules for *Ld*CysK - N-(2-{1-[(6-methoxy2,3-dihydro-1H-inden-5-yl)methyl]pyrrolidin-2-yl}-6-methylpyridin-4-yl)pyrazin2-amine (Lig1) and N-ethyl-N-{1-[(2-methyl-2,3-dihydro-1,3-benzothiazol-6yl)carbonyl]-1-azaspiro[4.5]decan-3-yl}acetamide (Lig2) via *in silico* screening of tetrapeptides with shape similarity to EWSI and DWSI (Kant et al., 2019). These ligands were docked into *Ld*CysK and Lig2 demonstrated improved interaction energy and capacity for hydrogen-bonding; this ligand was that which had greater conformational similarity to EWSI compared to DWSI. This potency of Lig2 was reinforced through molecular dynamics and binding energy analysis compared to Lig1 in complex with the enzyme, in terms of stability and compactness. In addition, both docking and molecular dynamic analyses demonstrated hydrogen bonding interactions between Lig2 and residues, Ser79 and Gln152 (Kant et al., 2019). Characterisation of these inhibitors both *in vitro* and *in vivo* remains to be investigated. It is worth noting that despite the significant advances that have been made in the chemical inhibitor space for CysK and CysM enzymes, the potencies of these inhibitors remain around 100-fold less effective than the complete enzymatic inhibition of CysK by CysE (Table 2.2).

2.9 Conclusion

There have been many campaigns aimed at discovering potent and selective inhibitors for the cysteine synthesis enzymes CysE and CysK. Yet despite excellent inhibition of activity seen for certain compounds many failed to inhibit bacterial growth, presumably due to the lack of permeability of the compounds across the bacterial wall. The best inhibitor for CysE that was bactericidal and had an IC_{50} of 48 µM, although this is much greater than the natural inhibitor cysteine, with an IC_{50} between 1 and 10 µM. Recently inhibitors of CysK were identified that demonstrate potent inhibition (nanomolar binding) and an adjuvant effect when used in combination with the antibiotic colistin. Based on this it is worth testing other promising inhibitors identified with known antibiotics to see if they have an adjuvant effect. Given that cysteine biosynthesis is often dispensable under nutrient

rich conditions but becomes more essential during infection and persistence it would also be worth testing if the identified compounds reduces infection and/or enhances clearance by the host immune system. As shown here there is increasing evidence for bacterial *de novo* cysteine biosynthesis as a promising drug target for either new antimicrobials or antibiotic adjuvants. Further validation of this pathway and further exploration of new and existing inhibitor compounds is vital to develop potent and selective inhibitors to overcome antimicrobial resistance in a range of Gram-positive and Gram-negative human pathogens.

2.10 References

Abd El-Aleam, R. H., George, R. F., Georgey, H. H., & Abdel-Rahman, H. M. (2021). Bacterial virulence factors: a target for heterocyclic compounds to combat bacterial resistance. *RSC Adv, 11*(58), 36459-36482. 10.1039/d1ra06238g

Agarwal, S. M., Jain, R., Bhattacharya, A., & Azam, A. (2008). Inhibitors of Escherichia coli serine acetyltransferase block proliferation of Entamoeba histolytica trophozoites. *International Journal for Parasitology, 38*(2), 137-141. 10.1016/j.ijpara.2007.09.009

Agren, D., Schnell, R., & Schneider, G. (2009). The C-terminal of CysM from Mycobacterium tuberculosis protects the aminoacrylate intermediate and is involved in sulphur donor selectivity. *FEBS Lett, 583*(2), 330-336. 10.1016/j.febslet.2008.12.019

Akerley, B. J., Rubin, E. J., Novick, V. L., Amaya, K., Judson, N., & Mekalanos, J. J. (2002). A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae. *Proc Natl Acad Sci U S A, 99*(2), 966-971. 10.1073/pnas.012602299

Alexeev, D., Baxter, R. L., Campopiano, D. J., Kerbarh, O., Sawyer, L., Tomczyk, N., . . . Webster, S. P. (2006). Suicide inhibition of α-oxamine synthases: structures of the covalent adducts of 8-amino-7-oxononanoate synthase with trifluoroalanine. *Organic & Biomolecular Chemistry, 4*(7), 1209-1212. 10.1039/B517922J

Alston, T. A., Muramatsu, H., Ueda, T., & Bright, H. J. (1981). Inactivation of gamma-cystathionase by gamma-fluorinated amino acids. *FEBS Lett, 128*(2), 293-297. 10.1016/00145793(81)80102-0

Amori, L., Katkevica, S., Bruno, A., Campanini, B., Felici, P., Mozzarelli, A., & Costantino, G. (2012). Design and synthesis of trans-2-substituted-cyclopropane-1-carboxylic acids as the first non-natural small molecule inhibitors of O-acetylserine sulfhydrylase. *MedChemComm, 3*(9), 1111-1116. 10.1039/C2MD20100C

Annunziato, G., Pieroni, M., Benoni, R., Campanini, B., Pertinhez, T. A., Pecchini, C., . . . Costantino, G. (2016). Cyclopropane-1,2-dicarboxylic acids as new tools for the biophysical investigation of O-acetylserine sulfhydrylases by fluorimetric methods and saturation transfer difference (STD) NMR. *Journal of Enzyme Inhibition and Medicinal Chemistry, 31*(sup4), 78-87. 10.1080/14756366.2016.1218486

Annunziato, G., Spadini, C., Franko, N., Storici, P., Demitri, N., Pieroni, M., Costantino, G. (2021). Investigational Studies on a Hit Compound Cyclopropane–Carboxylic Acid Derivative Targeting O-Acetylserine Sulfhydrylase as a Colistin Adjuvant. *ACS Infectious Diseases, 7*(2), 281-292. 10.1021/acsinfecdis.0c00378

Annunziato, G., Spadini, C., Marchetti, M., Franko, N., Pavone, M., Iannarelli, M., . . . Costantino, G. (2022). Inhibitors of O-Acetylserine Sulfhydrylase with a

Cyclopropane-Carboxylic Acid Scaffold Are Effective Colistin Adjuvants in Gram Negative Bacteria. *Pharmaceuticals (Basel), 15*(6) 10.3390/ph15060766

Azam, M. A., & Jayaram, U. (2016). Inhibitors of alanine racemase enzyme: a review. *J Enzyme Inhib Med Chem, 31*(4), 517-526. 10.3109/14756366.2015.1050010

Bachman, M. A., Breen, P., Deornellas, V., Mu, Q., Zhao, L., Wu, W., Mobley, H. L. (2015). Genome-Wide Identification of Klebsiella pneumoniae Fitness Genes during Lung Infection. *mBio, 6*(3), e00775. 10.1128/mBio.00775-15

Becker, D., Selbach, M., Rollenhagen, C., Ballmaier, M., Meyer, T. F., Mann, M., & Bumann, D. (2006). Robust Salmonella metabolism limits possibilities for new antimicrobials. *Nature, 440*(7082), 303-307. 10.1038/nature04616

Becker, M. A., & Tomkins, G. M. (1969). Pleiotropy in a Cysteine-requiring Mutant of Salmonella typhimurium Resulting from Altered Protein-Protein Interaction. *Journal of Biological Chemistry, 244*(21), 6023-6030. 10.1016/S0021-9258(18)63576-8

Benoni, R., Beck, C. M., Garza-Sánchez, F., Bettati, S., Mozzarelli, A., Hayes, C. S., & Campanini, B. (2017). Activation of an anti-bacterial toxin by the biosynthetic enzyme CysK: mechanism of binding, interaction specificity and competition with cysteine synthase. *Scientific Reports, 7*(1), 8817. 10.1038/s41598-017-09022-6

Benoni, R., De Bei, O., Paredi, G., Hayes, C. S., Franko, N., Mozzarelli, A., . . . Campanini, B. (2017). Modulation of Escherichia coli serine acetyltransferase catalytic activity in the cysteine synthase complex. *FEBS letters, 591*(9), 1212-1224. 10.1002/1873-3468.12630

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., & Duncan, K. (2002). Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. *Mol Microbiol, 43*(3), 717-731. 10.1046/j.1365-2958.2002.02779.x

Bhave, D. P., Muse, W., B III., & Carroll, K., S. . (2007). Drug Targets in Mycobacterial Sulphur Metabolism. *Infectious Disorders - Drug TargetsDisorders), 7*(2), 140-158. 10.2174/187152607781001772

Brunner, K., Maric, S., Reshma, R. S., Almqvist, H., Seashore-Ludlow, B., Gustavsson, A.-L., . . . Schneider, G. (2016). Inhibitors of the Cysteine Synthase CysM with Antibacterial Potency against Dormant Mycobacterium tuberculosis. *Journal of Medicinal Chemistry, 59*(14), 6848-6859. 10.1021/acs.jmedchem.6b00674

Brunner, K., Steiner, E. M., Reshma, R. S., Sriram, D., Schnell, R., & Schneider, G. (2017). Profiling of in vitro activities of urea-based inhibitors against cysteine synthases from Mycobacterium tuberculosis. *Bioorg Med Chem Lett, 27*(19), 4582-4587. 10.1016/j.bmcl.2017.08.039

Bruno, A., Amori, L., & Costantino, G. (2013). Computational Insights into the Mechanism of Inhibition of OASS-A by a Small Molecule Inhibitor. *Molecular Informatics, 32*(5-6), 447457. 10.1002/minf.201200174

Burkhard, P., Jagannatha Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F., & Jansonius, J. N. (1998). Three-dimensional structure of O-acetylserine sulfhydrylase from Salmonella typhimurium11Edited by R. Huber. *Journal of Molecular Biology, 283*(1), 121-133. 10.1006/jmbi.1998.2037

Burkhard, P., Tai, C. H., Ristroph, C. M., Cook, P. F., & Jansonius, J. N. (1999). Ligand binding induces a large conformational change in O-acetylserine sulfhydrylase from Salmonella typhimurium. *J Mol Biol, 291*(4), 941-953. 10.1006/jmbi.1999.3002

Burns-Huang, K., & Mundhra, S. (2019). Mycobacterium tuberculosis cysteine biosynthesis genes mec+-cysO-cysM confer resistance to clofazimine. *Tuberculosis, 115*, 63-66. 10.1016/j.tube.2019.02.002

Burns, K. E., Baumgart, S., Dorrestein, P. C., Zhai, H., McLafferty, F. W., & Begley, T. P. (2005). Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis. *J Am Chem Soc, 127*(33), 11602-11603. 10.1021/ja053476x

Campanini, B., Benoni, R., Bettati, S., Beck, C. M., Hayes, C. S., & Mozzarelli, A. (2015). Moonlighting O-acetylserine sulfhydrylase: New functions for an old protein. *Biochim Biophys Acta, 1854*(9), 1184-1193. 10.1016/j.bbapap.2015.02.013

Campanini, B., Pieroni, M., Raboni, S., Bettati, S., Benoni, R., Pecchini, C., . . . Mozzarelli, A. (2015). Inhibitors of the Sulphur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Current Medicinal Chemistry, 22*(2), 187-213. 10.2174/0929867321666141112122553

Campanini, B., Speroni, F., Salsi, E., Cook, P. F., Roderick, S. L., Huang, B., . . . Mozzarelli, A. (2005). Interaction of serine acetyltransferase with O-acetylserine sulfhydrylase active site: evidence from fluorescence spectroscopy. *Protein science : a publication of the Protein Society, 14*(8), 2115-2124. 10.1110/ps.051492805

Capel, E., Zomer, A. L., Nussbaumer, T., Bole, C., Izac, B., Frapy, E., . . . Coureuil, M. (2016). Comprehensive Identification of Meningococcal Genes and Small Noncoding RNAs Required for Host Cell Colonization. *mBio, 7*(4), e01173-01116. 10.1128/mBio.01173-16

Chattopadhyay, A., Meier, M., Ivaninskii, S., Burkhard, P., Speroni, F., Campanini, B., . . . Cook, P. F. (2007). Structure, Mechanism, and Conformational Dynamics of O-Acetylserine Sulfhydrylase from Salmonella typhimurium: Comparison of A and B Isozymes. *Biochemistry, 46*(28), 8315-8330. 10.1021/bi602603c

Chaudhuri, R. R., Allen, A. G., Owen, P. J., Shalom, G., Stone, K., Harrison, M., . . . Charles, I. G. (2009). Comprehensive identification of essential Staphylococcus aureus genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics, 10*, 291. 10.1186/1471-2164-10-291

Chen, C., Yan, Q., Tao, M., Shi, H., Han, X., Jia, L., . . . Ma, Y. (2019). Characterization of serine acetyltransferase (CysE) from methicillin-resistant Staphylococcus aureus and inhibitory effect of two natural products on CysE. *Microbial Pathogenesis, 131*, 218-226. 10.1016/j.micpath.2019.04.002

Claus, M. T., Zocher, G. E., Maier, T. H., & Schulz, G. E. (2005). Structure of the O-acetylserine sulfhydrylase isoenzyme CysM from Escherichia coli. *Biochemistry, 44*(24), 8620-8626. 10.1021/bi050485+

Colyer, T. E., & Kredich, N. M. (1996). In vitro characterization of constitutive CysB proteins from Salmonella typhimurium. *Molecular Microbiology, 21*(2), 247-256. 10.1046/j.13652958.1996.6301347.x

Faraci, W. S., & Walsh, C. T. (1989). Mechanism of inactivation of alanine racemase by .beta.,.beta.,.beta.-trifluoroalanine. *Biochemistry, 28*(2), 431-437. 10.1021/bi00428a004

Fasnacht, M., & Polacek, N. (2021). Oxidative Stress in Bacteria and the Central Dogma of Molecular Biology. *Front Mol Biosci, 8*, 671037. 10.3389/fmolb.2021.671037

Filutowicz, M., Wiater, A., & Hulanicka, D. (1982). Delayed Inducibility of Sulphite Reductase in cysM Mutants of Salmonella typhimurium Under Anaerobic Conditions. *Microbiology, 128*(8), 1791-1794. 10.1099/00221287-128-8-1791

Francois, J. A., Kumaran, S., & Jez, J. M. (2006). Structural basis for interaction of O-acetylserine sulfhydrylase and serine acetyltransferase in the Arabidopsis cysteine synthase complex. *Plant Cell, 18*(12), 3647-3655. 10.1105/tpc.106.047316

Franko, N., Grammatoglou, K., Campanini, B., Costantino, G., Jirgensons, A., & Mozzarelli, A. (2018). Inhibition of O-acetylserine sulfhydrylase by fluoroalanine derivatives. *J Enzyme Inhib Med Chem, 33*(1), 1343-1351. 10.1080/14756366.2018.1504040

Frávega, J., Álvarez, R., Díaz, F., Inostroza, O., Tejías, C., Rodas, P. I., . . . Gil, F. (2016). Salmonella Typhimurium exhibits fluoroquinolone resistance mediated by the

accumulation of the antioxidant molecule H2S in a CysK-dependent manner. *J Antimicrob Chemother, 71*(12), 3409-3415. 10.1093/jac/dkw311

Gallagher, L. A., Ramage, E., Jacobs, M. A., Kaul, R., Brittnacher, M., & Manoil, C. (2007). A comprehensive transposon mutant library of Francisella novicida, a bioweapon surrogate. *Proceedings of the National Academy of Sciences, 104*(3), 1009-1014. doi:10.1073/pnas.0606713104

Gislason, A. S., Turner, K., Domaratzki, M., & Cardona, S. T. (2017). Comparative analysis of the Burkholderia cenocepacia K56-2 essential genome reveals cell envelope functions that are uniquely required for survival in species of the genus Burkholderia. *Microb Genom, 3*(11) 10.1099/mgen.0.000140

Gorman, J., & Shapiro, L. (2004). Structure of serine acetyltransferase from Haemophilus influenzae Rd. *Acta Crystallogr D Biol Crystallogr, 60*(Pt 9), 1600-1605. 10.1107/s0907444904015240

Guédon, E., & Martin-Verstraete, I. (2007). Cysteine Metabolism and Its Regulation in Bacteria. In V. F. Wendisch (Ed.), *Amino Acid Biosynthesis ~ Pathways, Regulation and Metabolic Engineering* (pp. 195-218). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hampshire, T., Soneji, S., Bacon, J., James, B. W., Hinds, J., Laing, K., . . . Butcher, P. D. (2004). Stationary phase gene expression of Mycobacterium tuberculosis following a progressive nutrient depletion: a model for persistent organisms? *Tuberculosis (Edinb), 84*(3-4), 228238. 10.1016/j.tube.2003.12.010

Hicks, J. L., & Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology, 164*(12), 1471-1480. 10.1099/mic.0.000728

Hindson, V. J. (2003). Serine acetyltransferase of Escherichia coli: substrate specificity and feedback control by cysteine. *The Biochemical journal, 375*(Pt 3), 745-752. 10.1042/BJ20030429

Hindson, V. J., & Shaw, W. V. (2003). Random-Order Ternary Complex Reaction Mechanism of Serine Acetyltransferase from Escherichia coli. *Biochemistry, 42*(10), 3113-3119. 10.1021/bi0267893

Hryniewicz, M. M., & Kredich, N. M. (1994). Stoichiometry of binding of CysB to the cysJIH, cysK, and cysP promoter regions of Salmonella typhimurium. *J Bacteriol, 176*(12), 3673-3682. 10.1128/jb.176.12.3673-3682.1994

Huang, B., Vetting, M. W., & Roderick, S. L. (2005). The active site of O-acetylserine sulfhydrylase is the anchor point for bienzyme complex formation with serine acetyltransferase. *Journal of bacteriology, 187*(9), 3201-3205. 10.1128/JB.187.9.3201-3205.2005

Jean Kumar, V. U., Poyraz, Ö., Saxena, S., Schnell, R., Yogeeswari, P., Schneider, G., & Sriram, D. (2013). Discovery of novel inhibitors targeting the Mycobacterium tuberculosis Oacetylserine sulfhydrylase (CysK1) using virtual high-throughput screening. *Bioorganic & Medicinal Chemistry Letters, 23*(5), 1182-1186. 10.1016/j.bmcl.2013.01.031

John, R. A., & Tudball, N. (1972). Evidence for induced fit of a pseudo-substrate of aspartate aminotransferase. *Eur J Biochem, 31*(1), 135-138. 10.1111/j.1432-1033.1972.tb02510.x

Johnson, C. M., Huang, B., Roderick, S. L., & Cook, P. F. (2004). Kinetic mechanism of the serine acetyltransferase from Haemophilus influenzae. *Archives of Biochemistry and Biophysics, 429*(2), 115-122. 10.1016/j.abb.2004.06.006

Joshi, P., Gupta, A., & Gupta, V. (2019). Insights into multifaceted activities of CysK for therapeutic interventions. *3 Biotech, 9*(2), 44. 10.1007/s13205-019-1572-4

Jovanovic, M., Lilic, M., Savic, D. J., & Jovanovic, G. (2003). The LysR-type transcriptional regulator CysB controls the repression of hslJ transcription in Escherichia coli. *Microbiology (Reading), 149*(Pt 12), 3449-3459. 10.1099/mic.0.26609-0

Kant, V., Vijayakumar, S., Sahoo, G. C., Ali, V., Singh, K., Chaudhery, S. S., & Das, P. (2019). In silico screening and validation of high-affinity tetra-peptide inhibitor of Leishmania donovani O-acetyl serine sulfhydrylase (OASS). *J Biomol Struct Dyn, 37*(2), 481-492. 10.1080/07391102.2018.1429315

Kesicki, E. A., Bailey, M. A., Ovechkina, Y., Early, J. V., Alling, T., Bowman, J., . . . Parish, T. (2016). Synthesis and Evaluation of the 2-Aminothiazoles as Anti-Tubercular Agents. *PLOS ONE, 11*(5), e0155209. 10.1371/journal.pone.0155209

Kim, Y., Zhou, M., Peterson, S., Anderson, W. F., & Joachimiak, A. (2006). *Crystal Structure of Serine Acetyltransferase CysE from Yersinia pestis*.

Kredich, N. (2008). Biosynthesis of Cysteine. *EcoSal Plus* doi:10.1128/ecosalplus.3.6.1.11

Kredich, N. M. (1992). The molecular basis for positive regulation of cys promoters in Salmonella typhimurium and Escherichia coli. *Mol Microbiol, 6*(19), 2747-2753. 10.1111/j.13652958.1992.tb01453.x

Kredich, N. M., Becker, M. A., & Tomkins, G. M. (1969). Purification and Characterization of Cysteine Synthetase, a Bifunctional Protein Complex, from Salmonella typhimurium. *Journal of Biological Chemistry, 244*(9), 2428-2439.

Kredich, N. M., & Tomkins, G. M. (1966). The Enzymic Synthesis of l-Cysteine in Escherichia coli and Salmonella typhimurium. *Journal of Biological Chemistry, 241*(21), 4955-4965.

Kumar, S., Kumar, N., Alam, N., & Gourinath, S. (2014). Crystal structure of serine acetyl transferase from Brucella abortus and its complex with coenzyme A. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1844*(10), 1741-1748. 10.1016/j.bbapap.2014.07.009

Kumar, S., Raj, I., Nagpal, I., Subbarao, N., & Gourinath, S. (2011). Structural and Biochemical Studies of Serine Acetyltransferase Reveal Why the Parasite Entamoeba histolytica Cannot Form a Cysteine Synthase Complex. *Journal of Biological Chemistry, 286*(14), 1253312541. 10.1074/jbc.M110.197376

Le Faou, A. (1984). Sulphur nutrition and metabolism in various species of Neisseria. *Annales de l'Institut Pasteur / Microbiologie, 135*(1, Supplement B), 3-11. 10.1016/S07692609(84)80037-X

Liang, J., Han, Q., Tan, Y., Ding, H., & Li, J. (2019). Current Advances on Structure-Function Relationships of Pyridoxal 5′-Phosphate-Dependent Enzymes. *Frontiers in Molecular Biosciences, 6* 10.3389/fmolb.2019.00004

Lithgow, J. K., Hayhurst, E. J., Cohen, G., Aharonowitz, Y., & Foster, S. J. (2004). Role of a cysteine synthase in Staphylococcus aureus. *Journal of bacteriology, 186*(6), 1579-1590. 10.1128/jb.186.6.1579-1590.2004

Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D., & Hryniewicz, M. M. (2001). Functional dissection of the LysR-type CysB transcriptional regulator. Regions important for DNA binding, inducer response, oligomerization, and positive control. *J Biol Chem, 276*(3), 2098-2107. 10.1074/jbc.M007192200

Magalhães, J., Franko, N., Raboni, S., Annunziato, G., Tammela, P., Bruno, A., . . . Costantino, G. (2021). Discovery of Substituted (2-Aminooxazol-4-yl)Isoxazole-3-carboxylic Acids as Inhibitors of Bacterial Serine Acetyltransferase in the Quest for Novel Potential Antibacterial Adjuvants. *Pharmaceuticals, 14*(2), 174.

Magalhães, J., Franko, N., Raboni, S., Annunziato, G., Tammela, P., Bruno, A., . . . Costantino, G. (2020). Inhibition of Nonessential Bacterial Targets: Discovery of a Novel Serine O-Acetyltransferase Inhibitor. *ACS Medicinal Chemistry Letters* 10.1021/acsmedchemlett.9b00627

Mandal, R. K., Jiang, T., & Kwon, Y. M. (2017). Essential genome of Campylobacter jejuni. *BMC Genomics, 18*(1), 616. 10.1186/s12864-017-4032-8

Marchetti, M., De Angelis, F. S., Annunziato, G., Costantino, G., Pieroni, M., Ronda, L., . . . Bettati, S. (2021). A Competitive O-Acetylserine Sulfhydrylase Inhibitor Modulates the Formation of Cysteine Synthase Complex. *Catalysts, 11*(6), 700.

Meissner, A., Boshoff, H. I., Vasan, M., Duckworth, B. P., Barry, C. E., & Aldrich, C. C. (2013). Structure–activity relationships of 2-aminothiazoles effective against Mycobacterium tuberculosis. *Bioorganic & Medicinal Chemistry, 21*(21), 6385-6397. 10.1016/j.bmc.2013.08.048

Mino, K., Hiraoka, K., Imamura, K., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2000). Characteristics of serine acetyltransferase from Escherichia coli deleting different lengths of amino acid residues from the C-terminus. *Biosci Biotechnol Biochem, 64*(9), 1874-1880. 10.1271/bbb.64.1874

Mino, K., & Ishikawa, K. (2003). A novel O-phospho-L-serine sulfhydrylation reaction catalysed by O-acetylserine sulfhydrylase from Aeropyrum pernix K1. *FEBS Lett, 551*(1-3), 133-138. 10.1016/s0014-5793(03)00913-x

Mino, K., Yamanoue, T., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2000). Effects of bienzyme complex formation of cysteine synthetase from escherichia coli on some properties and kinetics. *Biosci Biotechnol Biochem, 64*(8), 1628-1640. 10.1271/bbb.64.1628

Momitani, K., Shiba, T., Sawa, T., Ono, K., & Hurukawa, S. (2022). *Crystal structure of serine acetyltransferase from Salmonella typhimurium*.

Moule, M. G., Hemsley, C. M., Seet, Q., Guerra-Assunção, J. A., Lim, J., Sarkar-Tyson, M., . . . Wren, B. W. (2014). Genome-wide saturation mutagenesis of Burkholderia pseudomallei K96243 predicts essential genes and novel targets for antimicrobial development. *mBio, 5*(1), e00926-00913. 10.1128/mBio.00926-13

Muir, A., Gurung, I., Cehovin, A., Bazin, A., Vallenet, D., & Pelicic, V. (2020). Construction of a complete set of Neisseria meningitidis mutants and its use for the phenotypic profiling of this human pathogen. *Nature Communications, 11*(1), 5541. 10.1038/s41467-020-19347-y

Neuwald, A. F., Krishnan, B. R., Brikun, I., Kulakauskas, S., Suziedelis, K., Tomcsanyi, T., . . . Berg, D. E. (1992). cysQ, a gene needed for cysteine synthesis in Escherichia coli K-12 only during aerobic growth. *J Bacteriol, 174*(2), 415-425. 10.1128/jb.174.2.415-425.1992

Noji, M., Takagi, Y., Kimura, N., Inoue, K., Saito, M., Horikoshi, M., . . . Saito, K. (2001). Serine Acetyltransferase Involved in Cysteine Biosynthesis from Spinach: Molecular Cloning, Characterization and Expression Analysis of cDNA Encoding a Plastidic Isoform. *Plant and Cell Physiology, 42*(6), 627-634. 10.1093/pcp/pce078

Oldham, K. E. A., Prentice, E. J., Summers, E. L., & Hicks, J. L. (2022). Serine acetyltransferase from Neisseria gonorrhoeae; structural and biochemical basis of inhibition. *Biochem J, 479*(1), 57-74. 10.1042/bcj20210564

Olsen, L., B., H., Vetting, M., & Roderick, S. (2004). Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor. *Biochemistry, 43*(20), 6013-6019.

Oppezzo, O. J., & Antón, D. N. (1995). Involvement of cysB and cysE genes in the sensitivity of Salmonella typhimurium to mecillinam. *Journal of bacteriology, 177*(15), 4524-4527. 10.1128/jb.177.15.4524-4527.1995

Park, S., & Imlay, J. A. (2003). High Levels of Intracellular Cysteine Promote Oxidative DNA Damage by Driving the Fenton Reaction. *Journal of Bacteriology, 185*(6), 1942-1950. 10.1128/jb.185.6.1942-1950.2003

Passera, E., Campanini, B., Rossi, F., Casazza, V., Rizzi, M., Pellicciari, R., & Mozzarelli, A. (2011). Human kynurenine aminotransferase II – reactivity with substrates and inhibitors. *The FEBS Journal, 278*(11), 1882-1900. 10.1111/j.1742-4658.2011.08106.x

Pawar, A., Jha, P., Chopra, M., Chaudhry, U., & Saluja, D. (2020). Screening of natural compounds that targets glutamate racemase of Mycobacterium tuberculosis reveals

the anti-tubercular potential of flavonoids. *Scientific Reports, 10*(1), 949. 10.1038/s41598-020-57658-8

Phillips, R. S., & Dua, R. K. (1992). Indole protects tryptophan indole-lyase, but not tryptophan synthase, from inactivation by trifluoroalanine. *Arch Biochem Biophys, 296*(2), 489-496. 10.1016/0003-9861(92)90602-s

Pieroni, M., Annunziato, G., Beato, C., Wouters, R., Benoni, R., Campanini, B., . . . Costantino, G. (2016). Rational Design, Synthesis, and Preliminary Structure–Activity Relationships of αSubstituted-2-Phenylcyclopropane Carboxylic Acids as Inhibitors of Salmonella typhimurium O-Acetylserine Sulfhydrylase. *Journal of Medicinal Chemistry, 59*(6), 25672578. 10.1021/acs.jmedchem.5b01775

Poulin, R., Lu, L., Ackermann, B., Bey, P., & Pegg, A. E. (1992). Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by alpha-difluoromethylornithine. Characterization of sequences at the inhibitor and coenzyme binding sites. *J Biol Chem, 267*(1), 150-158.

Poyraz, Ö., Jeankumar, V. U., Saxena, S., Schnell, R., Haraldsson, M., Yogeeswari, P., . . . Schneider, G. (2013). Structure-Guided Design of Novel Thiazolidine Inhibitors of O-Acetyl Serine Sulfhydrylase from Mycobacterium tuberculosis. *Journal of Medicinal Chemistry, 56*(16), 6457-6466. 10.1021/jm400710k

Pye, V. E., Tingey, A. P., Robson, R. L., & Moody, P. C. E. (2004). The Structure and Mechanism of Serine Acetyltransferase from Escherichia coli. *Journal of Biological Chemistry, 279*(39), 40729-40736. 10.1074/jbc.M403751200

Rabeh, W. M., & Cook, P. F. (2004). Structure and Mechanism of O-Acetylserine Sulfhydrylase. *Journal of Biological Chemistry, 279*(26), 26803-26806. 10.1074/jbc.R400001200

Raetz, C. R. H., & Roderick, S. L. (1995). A Left-Handed Parallel β Helix in the Structure of UDP-N-Acetylglucosamine Acyltransferase. *Science, 270*(5238), 997-1000. 10.1126/science.270.5238.997

Raj, I., Kumar, S., & Gourinath, S. (2012). The narrow active-site cleft of O-acetylserine sulfhydrylase from Leishmania donovani allows complex formation with serine acetyltransferases with a range of C-terminal sequences. *Acta Crystallogr D Biol Crystallogr, 68*(Pt 8), 909-919. 10.1107/s0907444912016459

Rakonjac, J., Milic, M., & Savic, D. J. (1991). cysB and cysE mutants of Escherichia coli K12 show increased resistance to novobiocin. *Mol Gen Genet, 228*(1-2), 307-311. 10.1007/bf00282481

Raman, K., Yeturu, K., & Chandra, N. (2008). targetTB: a target identification pipeline for Mycobacterium tuberculosis through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol, 2*, 109. 10.1186/1752-0509-2-109

Ramírez, A., Castañeda, M., Xiqui, M. L., Sosa, A., & Baca, B. E. (2006). Identification, cloning and characterization of cysK, the gene encoding O-acetylserine (thiol)-lyase from Azospirillum brasilense, which is involved in tellurite resistance. *FEMS Microbiology Letters, 261*(2), 272-279. 10.1111/j.1574-6968.2006.00369.x

Reichelt, A., & Martin, S. F. (2006). Synthesis and Properties of Cyclopropane-Derived Peptidomimetics. *Accounts of Chemical Research, 39*(7), 433-442. 10.1021/ar030255s

Remmele, C. W., Xian, Y., Albrecht, M., Faulstich, M., Fraunholz, M., Heinrichs, E., . . . Rudel, T. (2014). Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic acids research, 42*(16), 10579-10595. 10.1093/nar/gku762

Rengarajan, J., Bloom, B. R., & Rubin, E. J. (2005). Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages. *Proc Natl Acad Sci U S A, 102*(23), 8327-8332. 10.1073/pnas.0503272102

Roop, R. M., 2nd, Gaines, J. M., Anderson, E. S., Caswell, C. C., & Martin, D. W. (2009). Survival of the fittest: how Brucella strains adapt to their intracellular niche in the host. *Med Microbiol Immunol, 198*(4), 221-238. 10.1007/s00430-009-0123-8

Rosa, B., Dickinson, E. R., Marchetti, M., Campanini, B., Pioselli, B., Bettati, S., & Rand, K. D. (2021). Revealing the Dynamic Allosteric Changes Required for Formation of the Cysteine Synthase Complex by Hydrogen-Deuterium Exchange MS. *Mol Cell Proteomics, 20*, 100098. 10.1016/j.mcpro.2021.100098

Rosa, B., Marchetti, M., Paredi, G., Amenitsch, H., Franko, N., Benoni, R., . . . Bettati, S. (2019). Combination of SAXS and Protein Painting Discloses the Three-Dimensional Organization of the Bacterial Cysteine Synthase Complex, a Potential Target for Enhancers of Antibiotic Action. *International Journal of Molecular Sciences, 20*(20), 5219.

Salsi, E., Bayden, A. S., Spyrakis, F., Amadasi, A., Campanini, B., Bettati, S., . . . Mozzarelli, A. (2010). Design of O-acetylserine sulfhydrylase inhibitors by mimicking nature. *Journal of medicinal chemistry, 53*(1), 345-356. 10.1021/jm901325e

Sassetti, C. M., & Rubin, E. J. (2003). Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A, 100*(22), 12989-12994. 10.1073/pnas.2134250100

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., . . . Schoolnik, G. K. (2003). Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. *The Journal of experimental medicine, 198*(5), 693-704. 10.1084/jem.20030846

Schnell, R., Oehlmann, W., Singh, M., & Schneider, G. (2007). Structural Insights into Catalysis and Inhibition of O-Acetylserine Sulfhydrylase from Mycobacterium tuberculosis: Crystal Structures of the Enzyme α-Aminoacrylate Intermediate and an Enzyme-Inhibitor Complex. *Journal of Biological Chemistry, 282*(32), 23473-23481. 10.1074/jbc.M703518200

Shatalin, K., Shatalina, E., Mironov, A., & Nudler, E. (2011). H2S: a universal defense against antibiotics in bacteria. *Science, 334*(6058), 986-990. 10.1126/science.1209855

Shukla, H., Kumar, V., Singh, A. K., Rastogi, S., Khan, S. R., Siddiqi, M. I., . . . Akhtar, M. S. (2015). Isocitrate lyase of Mycobacterium tuberculosis is inhibited by quercetin through binding at N-terminus. *International Journal of Biological Macromolecules, 78*, 137-141. 10.1016/j.ijbiomac.2015.04.005

Silverman, R. B. (1995). Mechanism-based enzyme inactivators. *Methods Enzymol, 249*, 240-283. 10.1016/0076-6879(95)49038-8

Silverman, R. B., & Abeles, R. H. (1976). Inactivation of pyridoxal phosphate dependent enzymes by mono- and polyhaloalanines. *Biochemistry, 15*(21), 4718-4723. 10.1021/bi00666a028

Silverman, R. B., & Abeles, R. H. (1977). Mechanism of inactivation of gamma-cystathionase by beta,beta,beta-trifluoroalanine. *Biochemistry, 16*(25), 5515-5520. 10.1021/bi00644a019

Singh, P., Brooks, J. F., 2nd, Ray, V. A., Mandel, M. J., & Visick, K. L. (2015). CysK Plays a Role in Biofilm Formation and Colonization by Vibrio fischeri. *Appl Environ Microbiol, 81*(15), 5223-5234. 10.1128/aem.00157-15

Spyrakis, F., Felici, P., Bayden, A. S., Salsi, E., Miggiano, R., Kellogg, G. E., . . . Campanini, B. (2013). Fine tuning of the active site modulates specificity in the interaction of O-acetylserine sulfhydrylase isozymes with serine acetyltransferase. *Biochim Biophys Acta, 1834*(1), 169-181. 10.1016/j.bbapap.2012.09.009

Spyrakis, F., Singh, R., Cozzini, P., Campanini, B., Salsi, E., Felici, P., . . . Mozzarelli, A. (2013). Isozyme-Specific Ligands for O-acetylserine sulfhydrylase, a Novel Antibiotic Target. *PLOS ONE, 8*(10), e77558. 10.1371/journal.pone.0077558

Stipanuk, M. H., Dominy, J. E., Jr., Lee, J. I., & Coloso, R. M. (2006). Mammalian cysteine metabolism: new insights into regulation of cysteine metabolism. *J Nutr, 136*(6 Suppl), 1652s-1659s. 10.1093/jn/136.6.1652S

Sturgill, G., Toutain, C. M., Komperda, J., O'Toole, G. A., & Rather, P. N. (2004). Role of CysE in Production of an Extracellular Signalling Molecule in Providencia stuartii and Escherichia coli: Loss of cysE Enhances Biofilm Formation in Escherichia coli. *Journal of Bacteriology, 186*(22), 7610-7617. doi:10.1128/jb.186.22.7610-7617.2004

Tai, C.-H., Burkhard, P., Gani, D., Jenn, T., Johnson, C., & Cook, P. F. (2001). Characterization of the Allosteric Anion-Binding Site of O-Acetylserine Sulfhydrylase. *Biochemistry, 40*(25), 7446-7452. 10.1021/bi015511s

Tai, C. H., Nalabolu, S. R., Jacobson, T. M., Minter, D. E., & Cook, P. F. (1993). Kinetic mechanisms of the A and B isozymes of O-acetylserine sulfhydrylase from Salmonella typhimurium LT-2 using the natural and alternate reactants. *Biochemistry, 32*(25), 64336442. 10.1021/bi00076a017

Takumi, K., & Nonaka, G. (2016). Bacterial Cysteine-Inducible Cysteine Resistance Systems. *J Bacteriol, 198*(9), 1384-1392. 10.1128/jb.01039-15

Toyomoto, T., Ono, K., Shiba, T., Momitani, K., Zhang, T., Tsutsuki, H., . . . Sawa, T. (2023). Alkyl gallates inhibit serine O-acetyltransferase in bacteria and enhance susceptibility of drug-resistant Gram-negative bacteria to antibiotics. *Frontiers in Microbiology, 14* 10.3389/fmicb.2023.1276447

Turnbull, A. L., & Surette, M. G. (2008). l-Cysteine is required for induced antibiotic resistance in actively swarming Salmonella enterica serovar Typhimurium. *Microbiology, 154*(11), 3410-3419. 10.1099/mic.0.2008/020347-0

Turnbull, A. L., & Surette, M. G. (2010). Cysteine biosynthesis, oxidative stress and antibiotic resistance in Salmonella typhimurium. *Research in Microbiology, 161*(8), 643-650. 10.1016/j.resmic.2010.06.004

Tysoe, C., & Withers, S. G. (2014). Fluorinated mechanism-based inhibitors: common themes and recent developments. *Curr Top Med Chem,14*(7), 865-874. 10.2174/1568026614666140202204602

Ulusu, N. N. (2015). Evolution of Enzyme Kinetic Mechanisms. *Journal of Molecular Evolution, 80*(5), 251-257. 10.1007/s00239-015-9681-0

Varesio, L. M., Fiebig, A., & Crosson, S. (2021). Brucella ovis Cysteine Biosynthesis Contributes to Peroxide Stress Survival and Fitness in the Intracellular Niche. *Infect Immun, 89*(6) 10.1128/iai.00808-20

Vásquez, C. C., Saavedra, C. P., Loyola, C. A., Araya, M. A., & Pichuantes, S. (2001). The Product of the cysK Gene of Bacillus stearothermophilus V Mediates Potassium Tellurite Resistance in Escherichia coli. *Current Microbiology, 43*(6), 418-423. 10.1007/s002840010331

Verma, D., Gupta, S., Saxena, R., Kaur, P., R, R., Srivastava, S., & Gupta, V. (2020). Allosteric inhibition and kinetic characterization of Klebsiella pneumoniae CysE: An emerging drug target. *Int J Biol Macromol, 151*, 1240-1249. 10.1016/j.ijbiomac.2019.10.170

Voskuil, M., Bartek, I., Visconti, K., & Schoolnik, G. (2011). The Response of Mycobacterium Tuberculosis to Reactive Oxygen and Nitrogen Species. *Frontiers in Microbiology, 2* 10.3389/fmicb.2011.00105

Voskuil, M. I., Visconti, K. C., & Schoolnik, G. K. (2004). Mycobacterium tuberculosis gene expression during adaptation to stationary phase and low-oxygen dormancy. *Tuberculosis (Edinb), 84*(3-4), 218-227. 10.1016/j.tube.2004.02.003

Wang, E. A., & Walsh, C. (1981). Characteristics of .beta.,.beta.-difluoroalanine and .beta.,.beta.,.beta.-trifluoroalanine as suicide substrates for E. coli B alanine racemase. *Biochemistry, 20*(26), 7539-7546. 10.1021/bi00529a032

Wang, T., & Leyh, T. S. (2012). Three-stage assembly of the cysteine synthase complex from Escherichia coli. *The Journal of biological chemistry, 287*(6), 4360-4367. 10.1074/jbc.M111.288423

Warr, A. R., Hubbard, T. P., Munera, D., Blondel, C. J., Abel zur Wiesch, P., Abel, S., . . . Waldor, M. K. (2019). Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization. *PLOS Pathogens, 15*(8), e1007652. 10.1371/journal.ppat.1007652

Wong, Y. C., Abd El Ghany, M., Naeem, R., Lee, K. W., Tan, Y. C., Pain, A., & Nathan, S. (2016). Candidate Essential Genes in Burkholderia cenocepacia J2315 Identified by Genome-Wide TraDIS. *Front Microbiol, 7*, 1288. 10.3389/fmicb.2016.01288

Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B., & Jez, J. M. (2013). Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. *The Journal of biological chemistry, 288*(51), 36463-36472. 10.1074/jbc.M113.527143

Zhao, C., Kumada, Y., Imanaka, H., Imamura, K., & Nakanishi, K. (2006). Cloning, overexpression, purification, and characterization of O-acetylserine sulfhydrylase-B from Escherichia coli. *Protein Expr Purif, 47*(2), 607-613. 10.1016/j.pep.2006.01.002

Zhao, C., Moriga, Y., Feng, B., Kumada, Y., Imanaka, H., Imamura, K., & Nakanishi, K. (2006). On the interaction site of serine acetyltransferase in the cysteine synthase complex from Escherichia coli. *Biochemical and Biophysical Research Communications, 341*(4), 911916. 10.1016/j.bbrc.2006.01.054

Chapter Three: Serine acetyltransferase from *Neisseria gonorrhoeae*; structural and biochemical basis of inhibition

3.1 Preface

De novo synthesis of the amino acid L-cysteine, is well-conserved across bacteria, protists and plants and is a primary mechanism for the acquisition and assimilation of inorganic sulphur sources. Scavenging of nutrients is well-established to play a key role in the success of a pathogen to adapt to the host environment. A previous review by our research group (Hicks & Mullholland, 2018) identified in *Neisseria gonorrhoeae* a genomic disruption present in the sulphate reduction arm of the cysteine biosynthesis pathway, but the key biosynthetic enzymes serine *O*-acetyltransferase (CysE/SAT) and *O*-acetylserine sulphydrylase (CysK/OASSA) remained intact. Reports of *N. gonorrhoeae* being able to grow on thiosulphate as a sole sulphur source (Le Faou, 1984) coupled with a transposon mutagenesis study showing that *cysE* is an essential gene for *N. gonorrhoeae* (Remmele et al., 2014), prompts further investigation into functionality and role of this enzyme in *N. gonorrhoeae*. In this chapter, serine *O*-acetyltransferase from *N. gonorrhoeae* (*Ng*CysE) is structurally and kinetically characterised, using X-ray crystallography, small-angle X-ray scattering (SAXS), and spectrophotometric activity assays. Note in this manuscript "*Ng*SAT" has been replaced with "*Ng*CysE" to ensure consistency with other chapters.

The work presented in this chapter is shown in a publication format, as a published research article with supplementary information available in the Appendix A.

Citation for this publication is as follows:

Oldham, K. E. A., Prentice, E. J., Summers, E. L., & Hicks, J. L. (2022). Serine acetyltransferase from *Neisseria gonorrhoeae*; structural and biochemical basis of inhibition. *The Biochemical journal, 479*(1), 57-74. doi:10.1042/bcj20210564

3.1.1 Author contributions

I led the experimental work for this research, as first author, which involved enzyme purification, method optimisation and collection of kinetic assay data; optimisation of crystallisation conditions and data collection, solving, analysis, refinement and deposition of *Ng*CysE structures; as well as figure preparation, writing and editing of manuscript. Erica Prentice contributed to the interpretation and analysis of the kinetic data, figure preparation, and writing and editing of the manuscript. Emma Summers performed structural analysis of the structures, generation of figures and contributed to manuscript writing and editing. Joanna Hicks conceptualised research, provided funding for project, cloning of *Ng*CysE, figure preparation, collection and processing of SAXS data, contributed to manuscript writing and editing, and prepared the manuscript for submission. The authorship contribution form can be found in the Appendix D.

Serine acetyltransferase from *Neisseria gonorrhoeae*; structural and biochemical basis of inhibition

Keely E.A. Oldham[1,2], Erica J. Prentice[1], Emma L. Summers[1], Joanna L. Hicks[2]

[1]*Te aka Mātuatua School of Science, University of Waikato, Hamilton, New Zealand*

[2]*Te Huataki Waiora School of Health, University of Waikato, Hamilton, New Zealand*

3.2 Abstract

Serine acetyltransferase (CysE) catalyses the first step in the two-step pathway to synthesise L-cysteine in bacteria and plants. CysE synthesises *O*-acetylserine from substrates L-serine and acetyl-coenzyme A and is a key enzyme for regulating cellular cysteine levels by feedback inhibition of L-cysteine, and its involvement in the cysteine synthase complex. We have performed extensive structural and kinetic characterisation of the CysE enzyme from the antibiotic-resistant pathogen *Neisseria gonorrhoeae*. Using X-ray crystallography, we have solved the structures of *Ng*CysE with the non-natural ligand, L-malate (present in the crystallisation screen) to 2.01 Å and with the natural substrate L-serine (2.80 Å) bound. Both structures are hexamers, with each monomer displaying the characteristic lefthanded parallel β-helix domain of the acyltransferase superfamily of enzymes. Each structure displays both extended and closed conformations of the C-terminal tail. L-malate bound in the active site results in an interesting mix of open and closed active site conformations, exhibiting a structural change mimicking the conformation of cysteine (inhibitor) bound structures from other organisms. Kinetic characterisation shows competitive inhibition of L-cysteine with substrates L-serine and acetyl-coenzyme A. The CysE reaction represents a key point for the regulation of cysteine biosynthesis and controlling cellular sulphur due to feedback inhibition by L-cysteine and formation of the cysteine synthase complex. Data presented here provide the structural and mechanistic basis for inhibitor design and given this enzyme is not present in humans could be explored to combat the rise of extensively antimicrobial resistant *N. gonorrhoeae*.

3.3 Introduction

Neisseria gonorrhoeae is an obligate human pathogen that causes the sexually transmitted infection gonorrhoea. With purported references to gonorrhoea in early texts including the bible (Leviticus 15:1-3) the bacterium *N. gonorrhoeae* has plagued humans for centuries. Of the many and varied treatments developed over decades, *N. gonorrhoeae* has developed resistance to nearly all of the antibiotics used for its treatment. As a consequence, we are now facing the reality of untreatable gonorrhoea with only one effective class of antibiotics, the cephalosporin's left to treat this common infection. With no vaccine available, gonorrhoea control relies on our ability to treat gonorrhoea infections and new antimicrobials are desperately needed for control of this disease in the future.

The failure to identify new targets for antibiotic treatment has been shown by the small number of antibiotics discovered in the last 50 years and can be attributed in part to a lack of knowledge of microbial metabolism inside the host (Campanini, Pieroni, et al., 2015). Targeting amino acid biosynthesis is a new and promising route for the development of new antimicrobials (Berney et al., 2015; Murima, McKinney, & Pethe, 2014). Synthesis of the amino acid cysteine is the primary pathway for the incorporation of sulphur into a variety of important cellular constituents including methionine, thiamin, biotin and co-enzyme A. Cysteine also plays important roles in protein molecules in Fe–S clusters, catalysis, protein folding and the formation of disulphide bonds. In microorganisms, cysteine provides protection from oxidative stress via reducing systems such as thioredoxin and glutathione.

In most bacteria, the cysteine biosynthesis pathway is well conserved. Sulphate is transported into the cell and via a series of enzymatic steps is successively reduced to sulphide via the sulphate reduction pathway, which is then incorporated into cysteine (Kredich & Tomkins, 1966). Under anaerobic conditions, thiosulphate is imported and *S*-sulphocysteine is produced which is then converted into cysteine by an uncharacterised process.

CysE catalyses the first step in the two-step reaction that makes L-cysteine from L-serine. The side chain hydroxyl of L-serine undergoes a co-enzyme A-dependent acetylation to form *O*-acetylserine (Pye et al., 2004). The second step is then

catalysed by one of two OASS isoforms; OASS-A/CysK which catalyses the βreplacement of acetate with sulphide to convert *O*-acetylserine to L-cysteine (Rabeh & Cook, 2004), or OASS-B/CysM which catalyses the β-replacement of acetate with thiosulphate to form *S*-sulphocysteine (Chattopadhyay et al., 2007), which is converted to L-cysteine via an unknown mechanism.

Bacteria generally have both OASS-A and OASS-B isoforms, thereby capable of growth on sulphate or thiosulphate as the sole source of sulphur. *Neisseria* species appear to have only one OASS isoform annotated as CysK (OASS-A) (Hicks & Mullholland, 2018), yet unlike other *Neisseria* species, *N. gonorrhoeae* cannot grow on sulphate as the sole source of sulphur (Le Faou, 1984) due to a 3.5 kb deletion in the sulphate reduction operon compared with other *Neisseria* species (Hicks & Mullholland, 2018). However, all *Neisseria* strains, including *N. gonorrhoeae*, are able to use thiosulphate as a sole source of sulphur (Le Faou, 1984), which suggests that functional CysE and OASS enzymes are present.

CysE and OASS form a bi-enzyme complex termed the cysteine synthase complex which is intricately involved in the control of sulphur metabolism (Benoni, De Bei, et al., 2017; Kumaran, Yi, Krishnan, & Jez, 2009). Many bi-enzyme complexes channel substrates from one enzyme to another. However, the cysteine synthase complex is unusual in that the flexible C-terminal tail of CysE inserts into the active site of OASS competing with the binding of *O*-acetylserine and inhibiting OASS activity (Campanini et al., 2005; Mino et al., 1999). CysE activity is enhanced in the complex (Benoni, De Bei, et al., 2017) and produces the pathway intermediate *O*-acetylserine which at high concentrations promotes dissociation of the complex (Cook & Wedding, 1977) and subsequent consumption of *O*-acetylserine for cysteine synthesis. The OASS variant, OASS-B that utilises thiosulphate to produce *S*-sulphocysteine does not form a complex with the CysE enzyme (Zhao et al., 2006). Due to differences in the pathways for sulphate reduction between other bacterial pathogens and *N. gonorrhoeae*, respectively, it is unknown if the CysE and OASS enzymes from *N. gonorrhoeae* form the cysteine synthase complex or how this regulates sulphur metabolism.

Other mechanisms of regulation for the *de novo* biosynthesis of cysteine include feedback inhibition of CysE by L-cysteine, transcriptional regulation of sulphur

acquisition and cysteine biosynthetic genes by transcriptional regulator CysB regulating sulphur acquisition and changes in the quaternary structure of CysE (Hirai et al., 2003; Wirtz & Hell, 2006). Due to the unique nature of sulphur acquisition for cysteine biosynthesis in *N. gonorrhoeae*, the essential nature of CysE for growth in *N. gonorrhoeae* (*Ng*CysE) (Remmele et al., 2014) and the targeting of CysE in other pathogens (Agarwal et al., 2008; Chen, Han, et al., 2019; Magalhães et al., 2020), *Ng*CysE could represent a novel drug target for the treatment of antimicrobial resistant gonorrhoea.

CysE belongs to the family of *O*-acetyltransferases (EC 2.3.1.30) found in both plants and bacteria and there is good structural conservation between bacterial and plant CysE enzymes, even though there is a large degree of sequence divergence (Olsen et al., 2004; Yi et al., 2013). The *O*-acetyltransferase family is defined by the hexapeptide repeat sequence [LIV]-[GAED]-X_2-[STAV]-X, that gives rise to the unique left-handed β-helix (LβH), characteristic of this family (Raetz & Roderick, 1995).

CysE enzymes consist of two structural domains, an N-terminal, α-helical domain and a C-terminal, LβH domain. The C-terminal LβH domain is structured like a triangular prism, which fits with two other monomers to form a trimer (Johnson, Roderick, & Cook, 2005). The resulting trimer dimerises to form a hexamer through hydrophobic interactions at the N-terminal face of each trimer (Pye et al., 2004). CysE enzymes are unique as they are the only hexapeptide acyltransferase to adopt a hexameric conformation, instead of the more common trimeric arrangement (Johnson et al., 2005). This higher oligomeric state allows for formation of the bienzyme cysteine synthase complex and regulation of sulphur flux (Benoni, De Bei, et al., 2017; Mino et al., 1999). To date the only instance of a trimeric CysE has been identified in the human parasite *Entamoeba histolytica* (Kumar et al., 2011).

Members of the LβH family, including CysE have a conserved histidine in the active site (Johnson et al., 2005). The proposed mechanism for the acetyl-transfer reaction of CysE, is by general acid–base catalysis, where the histidine, alongside a conserved aspartate form a catalytic triad with the substrate L-serine (Johnson et al., 2005; Pye et al., 2004). In *Escherichia coli* this aspartate residue (Asp143) is

positioned in close proximity (2.8 Å) to the well-conserved histidine (His158), which is predicted to transform the histidine into a strong base (Pye et al., 2004). His158 activates the hydroxyl group of L-serine, for nucleophilic attack on the acetyl thioester of acetyl-CoA. The nucleophilic attack on acetyl-CoA forms a ternary complex between the enzyme, L-serine and acetyl-CoA, before the histidine behaves as a general acid and donates a hydrogen to the sulphur group of acetyl-CoA, leading to collapse of the complex and product release (Johnson et al., 2005; Pye et al., 2004).

Initial research hypothesised a ping-pong reaction mechanism for the CysE (Leu & Cook, 1994). However, a combination of equilibrium isotope exchange experiments (Hindson & Shaw, 2003) and Lineweaver–Burke plots (Benoni, De Bei, et al., 2017; Hindson & Shaw, 2003; Johnson et al., 2004b) overwhelmingly supports a sequential binding mechanism. The exact type of sequential mechanism reported for CysE varies, with a random sequential mechanism reported for CysE from *E. coli* (Hindson & Shaw, 2003) and an ordered binding mechanism reported for CysE from *Haemophilus influenzae* (Johnson et al., 2004b), where acetyl-CoA binds first, followed by L-serine.

The cysteine biosynthetic pathway is feedback inhibited by L-cysteine binding to CysE. L-cysteine competes directly with L-serine for the active site displaying competitive inhibition (Johnson et al., 2004b). Analysis of co-crystallised structures of CysE with L-cysteine, shows L-cysteine bound in the L-serine binding pocket located in the CysE active site (Olsen et al., 2004; Pye et al., 2004; Verma et al., 2020), supporting the observation of L-cysteine being a competitive inhibitor relative to L-serine (Johnson, Huang, Roderick, & Cook, 2004a; Kredich & Tomkins, 1966). This mechanism of inhibition appears to be well-conserved across the CysE family (Hindson, 2003; Johnson et al., 2004b). Interestingly, L-cysteine has also been reported to be a competitive inhibitor relative to acetyl-CoA (Johnson et al., 2004b). CysE structures with L-cysteine bound show conformational differences in the C-terminal tail compared with apo-CysE structures with the C-terminal tail folding against CysE to obscure the active site in L-cysteine-bound structures. This folded tail precludes acetyl-CoA from binding to the acetyl-CoA binding site (Yi et al., 2013) and prevents the unwanted acetylation of L-cysteine, given that it is isostructural to L-serine.

To design new antimicrobial inhibitors to target critical pathways such as cysteine biosynthesis in *N. gonorrhoeae* we need high quality, accurate three-dimensional structures of the target enzymes and a detailed understanding of enzymatic mechanism. To this end, we have determined the kinetic parameters and regulation of CysE from *N. gonorrhoeae* and the structure of this enzyme to 2.01 Å with the non-natural ligand, L-malate (from the crystallisation condition) bound in the active site and with the natural substrate L-serine bound to 2.8 Å.

3.4 Materials and Methods

3.4.1 Cloning of *Ng*CysE for expression in *Escherichia coli*

The CysE gene (annotated cysE) NGFG_RS07905 (old locus tag NGFG_01496) from *N. gonorrhoeae* MS11 was codon-optimised for *E. coli* and ordered from Geneart (Thermo Fisher). The synthetic NGFG_RS07905 construct was cloned into expression vector pET28b between Ndel and Xhol restriction sites for expression with an N-terminal His-tag. Insertion of NGFG_RS07905 into pET28b was confirmed by DNA sequencing prior to transformation into *E. coli* BL21 (DE3) for protein expression. Positive transformants were selected for by plating on Luria–Bertani (LB) agar supplemented with 50 $\mu g.ml^{-1}$ kanamycin and incubating overnight at 37°C.

3.4.2 *Ng*CysE expression and purification

E. coli BL21 (DE3) containing the NGFG_RS07905_pET28b plasmid were cultured in 1 L LB broth, supplemented with 50 $\mu g.ml^{-1}$ kanamycin. Cultures were incubated at 37°C (200 rpm) until OD_{600} reached between 0.5 and 0.7. Protein expression was induced by the addition of 0.75 mM IPTG and the cultures were incubated at 37°C (200 rpm) overnight. Cultures were centrifuged at 4600*xg* for 20 min at 4°C and the resulting cell pellet resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 20 mM imidazole). One Complete Mini, EDTA-free protease inhibitor tablet (Roche) was added prior to cell lysis by sonication. Lysate was centrifuged at 20,000*xg* for 20 min at room temperature and 0.2 μm filtered supernatant was loaded onto a pre-equilibrated HisTrap™ column (GE Healthcare). The column was washed with 20 ml of lysis buffer before the elution of *Ng*CysE using a 50% gradient over 25 ml (50 mM Tris pH 8.0, 200 mM NaCl, 1 M

imidazole). Immobilised-metal ion affinity chromatography-purified *Ng*CysE was concentrated at 15°C using a spin concentrator to a final volume of 0.75 ml. Concentrated *Ng*CysE was loaded and run through an Enrich 650 analytical gel filtration column (Bio-Rad), pre-equilibrated in 50 mM Tris pH 8.0, 100 mM NaCl and eluted *Ng*CysE was collected and stored at room temperature. Protein concentration was measured by absorbance at 280 nm by Nanodrop^{TM}.

3.4.3 Measuring the oligomeric state of *Ng*CysE

An Enrich 650 gel filtration column (Bio-Rad) was calibrated in 50 mM Tris pH 8.0, 100 mM NaCl using Gel Filtration Standards (Bio-Rad) according to manufacturer's instructions.

3.4.4 *Ng*CysE kinetic assays

*Ng*CysE for kinetic characterisation was purified immediately prior to assays. Enzyme was stored at room temperature for the duration of the assay, as a rapid decrease in activity was observed when stored on ice. Assays were conducted within 2 h post-purification as *Ng*CysE activity started to decrease after 2 h at room temperature (Figure A.1). *Ng*CysE activity was measured by adapting the method in (Noji, Inoue, Kimura, Gouda, & Saito, 1998). *Ng*CysE activity was measured by monitoring the decrease in absorbance at 232 nm (A_{232}) due to breakage of the thioester bond of acetyl-CoA ($\Delta\varepsilon 232 = 4500\ M^{-1}\ cm^{-1}$) using a Thermo Spectronic Heλios spectrophotometer (Thermo Fisher).

To calculate the K_M and k_{cat} for the substrate L-serine, assays were carried out in quartz cuvettes of 0.5 cm pathlength with a final reaction volume of 0.4 ml which contained 50 mM Tris pH 8.0, 100 mM NaCl, 0.45 mM acetyl-CoA and variable amounts of L-serine. The reaction was performed at 22°C with absorbance recorded every 0.125 s after the reaction was initiated by the addition of 0.625 mg of purified *Ng*CysE. Enzyme concentration in activity assays was optimised by testing various concentrations of *Ng*CysE (0.156, 0.312, 0.391, 0.521, 0.781 and 1.56 $mg.ml^{-1}$; Figure A.2). All substrate stocks were prepared in 50 mM Tris pH 8.0, 100 mM NaCl. An enzyme working stock of 0.125 $mg.ml^{-1}$ (3.96 mM, *Ng*CysE monomer 31.6 kDa) was stored at room temperature (22°C) for the duration of the assays. The K_M and k_{cat} was calculated for acetyl-CoA, by varying the amount of

acetyl-CoA and keeping the concentration of L-serine constant at 10 mM. K_M and k_{cat} values were determined by non-linear regression fit of the Michaelis–Menten equation (Equation 3.1) using GraphPad Prism (GraphPad Software Version 8.2.0). The initial velocity of the reaction was derived from linear-regression analysis of the first five seconds of the reaction. All concentrations were collected in duplicate. Rates were divided by enzyme concentration before plotting substrate concentration versus rate. Substrate inhibition was modelled for acetyl-CoA through non-linear regression fit of the substrate inhibition equation (Equation 3.2) using GraphPad Prism (GraphPad Software Version 8.2.0).

Equation 3.1 Michaelis-Menten equation.

$$Rate = \frac{V_{max}[S]}{K_M + [S]}$$

Equation 3.2 Substrate inhibition equation.

$$Rate = V_{max}\frac{[S]}{K_M + [S](1 + \frac{[S]}{K_i})}$$

3.4.5 *Ng*CysE kinetic assays

To calculate the IC_{50} for L-cysteine, similar reactions to above were set up with the addition of L-cysteine (final concentration, 0.01–40 μM) to the assay reaction before initiation of the reaction by the addition of *Ng*CysE. L-cysteine stocks (0.1 and 1.0 mM) were made fresh in 50 mM Tris pH 8.0, 100 mM NaCl, and stored on ice to prevent unwanted oxidation of L-cysteine to L-cystine. The IC_{50} for co-enzyme A (CoA) was calculated in a similar manner to cysteine, which involved the addition of CoA (final concentration, 10–2000 μM), before initiating the reaction with *Ng*CysE. The IC_{50} values for L-cysteine and CoA were determined by fitting Equation 3.3 using GraphPad Prism (GraphPad Software Version 8.2.0) (Gesztelyi et al., 2012; Weiss, 1997).

Equation 3.3 Log (inhibitor) versus normalised response - variable slope equation.
100

$$Rate = \frac{100}{(1 + 10^{((LogIC_{50} - X)Hillslope)})}$$

Michaelis–Menten plots for substrates acetyl-CoA and L-serine were collected as above with the addition of 4, 6 and 8 mM L-cysteine, respectively, to determine the mode of inhibition relative to each substrate. For varying concentrations of L-serine, acetyl-CoA was held constant at 0.15 mM and for varying concentrations of acetyl-CoA, L-serine was held constant at 1.5 mM. Inhibition data were analysed using the mixed-model inhibition equation (Equation 3.4) in GraphPad Prism (GraphPad Software Version 8.2.0) to determine the mode of inhibition for each substrate. K_M and k_{cat} values for each cysteine concentration relative to each substrate were determined by non-linear regression of the Michaelis–Menten equation (Equation 3.1) in GraphPad Prism (GraphPad Software Version 8.2.0).

Equation 3.4 Mixed-model inhibition equation.

$$Rate = V_{max}\frac{[S]}{[S]\left(1+\frac{[I]}{\alpha K_i}\right)+K_M(1+\frac{[I]}{K_i})}$$

3.4.6 Crystallisation of *Ng*CysE

Initial crystals were grown by vapor diffusion using an Index^{TM}-HR2-144 crystallisation screen (Hampton Research), dispensed into a low profile 96-2 well Intelli Plate (Hampton Research) using a Mosquito crystallisation robot (TTP LabTech Ltd). Sitting drops consisted of a 1:1 mix (100 nl :100 nl) of reservoir solution and concentrated protein (35 mg.ml^{-1} of purified *Ng*CysE) with 100 µl in the reservoir well. Crystals for diffraction collection were grown using hanging drop vapor diffusion at 18°C. *Ng*CysE crystals were obtained using whisker seeding in 28% (v/v) Tacsimate^{TM} pH 7.0 (0.3 M DL-malic acid, 1.8305 M malonic acid, 0.25 M ammonium citrate tribasic, 0.12 M succinic acid, 0.4 M sodium acetate trihydrate, 0.5 M sodium formate and 0.16 M ammonium tartrate dibasic, Hampton Research). Drop composition consisted of a 1:1 mix (2 µl:2 µl) of reservoir and protein with 500 µl in the reservoir well. *Ng*CysE with serine bound (*Ng*CysE + L-Ser) crystals were obtained using similar techniques, but were grown using 100 mM Tris pH 8.4, 26% (v/v) PEG400 and 15 mM serine (24 mg.ml^{-1} of purified *Ng*CysE).

For data collection, all crystals were transferred to a cryoprotectant solution, consisting of crystallisation solution with 15% (v/v) glycerol, prior to flash cooling in liquid nitrogen.

3.4.7 Data collection, indexing, integration, scaling

X-ray diffraction data (100 K) was collected at the Australian Synchrotron at the MX2 beamline (Aragao et al., 2018) equipped with an EIGER ×16M detector (Dectris). The *Ng*CysE and *Ng*CysE + L-Ser dataset diffraction images were indexed, integrated and scaled, using XDS (Kabsch, 2010). Merging of reflections was carried out using AIMLESS (Evans & Murshudov, 2013) from the CCP4 suite (Winn et al., 2011). Data quality was assessed through AIMLESS (Evans & Murshudov, 2013) and through manually viewing diffraction images using ALBULA (Dectris). The total number of monomers in the asymmetric unit was determined by calculating the solvent content using the Matthew's coefficient program (Matthews, 1968), as a part of the CCP4 program suite (Winn et al., 2011). The R free flag dataset was generated in AIMLESS (Evans & Murshudov, 2013). Data were analysed for evidence of twinning and translational non-crystallographic symmetry (tNCS) in AIMLESS from the CCP4 suite (Evans & Murshudov, 2013; Winn et al., 2011) and phenix.xtriage from the Phenix suite (Adams et al., 2010).

3.4.8 Structure building and refinement

The *Ng*CysE structure was solved using molecular replacement, using the structure of CysE from *Yersinia pestis* (3GVD) retrieved from the protein data bank (PDB). A single monomer was extracted from this file, in PyMOL (The PyMOL Molecular Graphics System, Version 2.3.2 Schrödinger) and used as the search model for molecular replacement using phenix.phaser (McCoy et al., 2007), from the Phenix suite (Adams et al., 2010). The *Ng*CysE + L-Ser structure was solved in a similar manner but using an *Ng*CysE (6WYE) monomer as the search model.

The models were initially built using the program phenix.autobuild (Terwilliger et al., 2008) from the Phenix suite (Adams et al., 2010). The resulting structure was manually built and refined using COOT (Emsley & Cowtan, 2004). For manual building, the $2F_o–F_c$ and $F_o– F_c$ electron density maps were set to 1σ and 3σ, respectively. Mobile sidechains without supported density were removed from

models. Translation-liberation-screw (TLS) was carried out in the late stages of model refinement using TLS groups determined by phenix.refine. After each round of structure manipulation, phenix.refine was used to run rounds of real-space refinement (Afonine et al., 2012).

3.4.9 Ligand fitting and refinement

The structures were prepared for ligand fitting by building and refining the protein structure and adding in waters using phenix.refine. Phenix.ligandfit was used to search for ligands of interest (Terwilliger, Klei, Adams, Moriarty, & Cohn, 2006; Venkatachalam, Jiang, Oldfield, & Waldman, 2003). Visual inspection and manual real-space refinement were carried out in COOT (Emsley & Cowtan, 2004). Metal ions were validated using the CheckMyMetal server (Zheng et al., 2017).

3.4.10 Structural analysis

Final structure statistics were generated using phenix.table_one, from the Phenix suite (Adams et al., 2010). Structure images were generated using PyMOL (Schrödinger, LLC). The final co-ordinate and structure amplitude files for *Ng*CysE and *Ng*CysE + L-Ser were deposited to the PDB, under accession codes 6WYE and 7RA4, respectively.

3.4.11 SAXS data collection

Measurements were performed at the Australian Synchrotron SAXS/WAXS beamline equipped with a Dectris-Pilatus detector. The wavelength of the X-rays was 1.0322 Å. The sample detector distance was 1426 mm. Data were collected from samples in a 1.5 mm thin-walled glass capillary at 25°C at two second intervals.

Size exclusion chromatography (SEC) in-line with the SAXS beamline (SECSAXS) (Ryan et al., 2018) with a co-flow set-up (Kirby et al., 2016) was used to collect scattering data. Data were collected from *Ng*CysE and *Ng*CysE + 50 mM L-cysteine following elution from a size exclusion column (Superdex 200 5/150), pre-equilibrated with SEC buffer (50 mM Tris pH 8.0, 100 mM NaCl) or SEC buffer with L-cysteine (50 mM Tris pH 8.0, 100 mM NaCl, 50 µM L-cysteine).

Raw scattering data were processed with Scatterbrain at the Australian Synchrotron. Scattered intensity (I) was plotted versus s using Primusqt from the ATSAS suite (Franke et al., 2017). All samples were devoid of an increase in intensity at low s (indicative of aggregation). Guinier plots were linear for sRg < 1.15. Theoretical scattering curves were generated from the *Ng*CysE crystal structure PDB file (6WYE) using Crysol (Svergun, Barberato, & Koch, 1995).

3.5 Results and Discussion

3.5.1 Purification and stoichiometry of *Ng*CysE

The *N. gonorrhoeae* serine acetyltransferase (CysE) gene NGFG_RS07905 was cloned into pET28b for expression with an N-terminal His-tag. This was done to avoid adding residues onto the C-terminus of the protein which is proposed to interact with OASS/CysK in order to form the cysteine synthase complex in later studies. *Ng*CysE was purified by IMAC followed by a final purification step via SEC. *Ng*CysE eluted as a single peak from a size exclusion column with an elution volume of 12.56 ml corresponding to an approximate molecular mass of 193.8 kDa (Figure A.3). The predicted molecular mass of the *Ng*CysE monomer is 31.6 kDa, confirming a hexamer of *Ng*CysE monomers (6 × 31.6 kDa = 189.6 kDa). SDS-PAGE analysis of *Ng*CysE shows >95% purity (Figure A.3).

3.5.2 Kinetic properties of *Ng*CysE

The activity of *Ng*CysE was determined by measuring the decrease in absorbance at 232 nm due to depletion of substrate acetyl-CoA. The kinetic parameters for substrate acetyl-CoA (Table 2.1) were calculated from fitting a Michaelis–Menten curve (R^2 = 0.8449) of rate (s^{-1}; rate divided by enzyme concentration) versus acetyl-CoA (black line, Figure 3.1A). The overall fit for the Michaelis–Menten equation is reasonable with the exception of *Ng*CysE rates collected at 1 mM acetyl-CoA which display inhibition that is likely to be substrate inhibition. We could not confidently fit the substrate inhibition equation (red line, Figure 3.1A) due to only one data point (1 mM acetyl-CoA) displaying inhibition. Due to limitations in experimental set-up (a high starting absorbance at high acetyl-CoA concentrations) we were unable to collect *Ng*CysE rates for higher acetyl-CoA concentrations.

However, data collected is highly reproducible with different enzyme preparations and enzyme concentrations (Figure A.4). From these data, the K_M (acetyl-CoA) was calculated to be 0.149 mM, and a k_{cat} of 1176 s^{-1} for the hexamer (Table 3.1).

Table 3.1 Kinetics parameters of *Ng*CysE.

Parameter	**L-serine[a]**	**Acetyl-CoA[a]**
K_M **(mM)**	1.21 ± 0.16	0.149 ± 0.05
k_{cat} **(s^{-1})[b]**	1444 ± 41	1176 ± 111
k_{cat}/K_M **($M^{-1}.s^{-1}$)**	1.19 x 10^6 ± 0.16	7.89 x 10^6 ± 2.54

[a]*Error is SEM of two replicates.*
[b]k_{cat} *calculated by dividing the rate by enzyme concentration using the concentration of the NgCysE hexamer (189.6 kDa).*

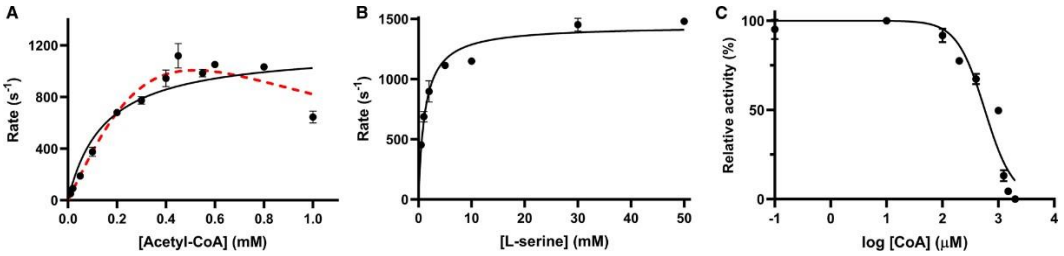


Figure 3.1 Kinetic analysis of *Ng*CysE substrates L-serine and acetyl-CoA. (A) Substrate inhibition is observed for acetyl-CoA. Michaelis–Menten (black) and substrate inhibition equations (dashed red lines), are modelled. (B) Michaelis- Menten fit for L-serine. L-serine and acetyl-CoA assays were collected at saturating concentrations of 0.45 mM acetyl-CoA and 10 mM L-serine, respectively. Enzyme concentration was factored into the rates in (A) and (B) by dividing by the enzyme concentration to give the rate (s^{-1}). (C) Dose response curve for co-enzyme A (CoA). Line represents the fit of IC_{50} equation to data. Substrate concentrations were fixed at 0.15 mM and 1.5 mM, for acetyl-CoA and L-serine, respectively. Plotted data points represent mean alongside SEM of two replicates.

A concentration of 0.45 mM acetyl-CoA was the saturating concentration used for collection of the L-serine Michaelis–Menten plot as this from this point the rate plateaus (Figure 3.1A) and higher acetyl-CoA concentrations (∼1 mM) display likely substrate inhibition. The kinetic parameters for substrate L-serine were calculated from the Michaelis–Menten curve (R^2 = 0.9476) of rate (s^{-1}; rate divided by enzyme concentration) versus L-serine concentration (Figure 3.1B). The K_M (Ser) of the enzyme was calculated to be 1.21 mM, the k_{cat} 1444 s^{-1} for the hexamer (Table 3.1).

To investigate if *Ng*CysE is inhibited by the high concentrations of the reaction product CoA as observed for other CysE enzymes (Hindson & Shaw, 2003)

(Johnson et al., 2004b) we collected an IC_{50} inhibition curve for CoA. There is inhibition of *Ng*CysE activity with increasing concentrations of CoA (Figure 3.1C) indicating product inhibition at high CoA concentrations with an IC_{50} value of 573 ± 11 µM, and a hillslope of −1.7 ± 0.3 indicating positive cooperativity. While CoA exhibits product inhibition, it does not account for the decrease in rate seen at 1 mM acetyl-CoA (Figure 3.1A) as reaction rates were calculated from steady-state conditions.

The difference in K_M values for each substrate (1.21 mM L-serine versus 0.149 mM acetyl-CoA), demonstrates a greater affinity of *Ng*CysE for acetyl-CoA compared with L-serine, which is consistent with other CysE homologues (Hindson & Shaw, 2003; Johnson et al., 2004b; Pye et al., 2004). The specificity constants, k_{cat}/K_M, for *Ng*CysE are 1.19×10^6 and 7.89×10^6 $M^{-1}.s^{-1}$ for L-serine and acetyl-CoA, respectively, which are below the diffusion theory maximum rate of ~10^8 $M^{-1}.s^{-1}$, indicating that the enzyme rate is not diffusion limited (Schurr & Schmitz, 1976).

There is a 6.5-fold increase in k_{cat}/K_M for acetyl-CoA compared with L-serine attributable to *Ng*CysE having a greater affinity for acetyl-CoA, with an approximately eight-fold lower K_M for acetyl-CoA (0.149 mM) compared with L-serine (1.21 mM). CysE enzymes from *E. coli* and *H. influenzae* display Michaelis– Menten kinetics for acetyl-CoA, but exhibit substrate inhibition for L-serine (Benoni, De Bei, et al., 2017; Johnson et al., 2004b; Kredich & Tomkins, 1966). *Ng*CysE appears to be the opposite, where acetyl-CoA possibly exhibits substrate inhibition and L-serine does not. Based on the reaction mechanism of CysE homologues (Benoni, De Bei, et al., 2017; Johnson et al., 2004b), we predict that *Ng*CysE has a sequential mechanism, in keeping with the proposed formation of the ternary complex during acyl transfer to L-serine.

The acetyl-CoA binding site of *Ng*CysE, may prove to be a good target for inhibitors, and this is being explored in other hexapeptide transferases. Inhibitors for a related hexapeptide transferase termed GlmU targeted the acetyl-CoA binding site, instead of the active site (*Ng*CysE equivalent L-serine binding site) (Olsen, Vetting, & Roderick, 2007).

3.5.3 Structure of *Ng*CysE

The crystal structure of *Ng*CysE was solved by molecular replacement to 2.01 Å with L-malate (from the crystallisation condition) bound in the active site. The final model was refined to an R and free R of 0.183 and 0.222, respectively (Table 3.2). The asymmetric unit contained six monomers arranged as one complete trimer and another trimer split into 1/3 and 2/3. A complete hexameric molecule is apparent after repetition by a crystallographic two-fold axis (Figure 3.2A). The most complete subunits contained residues 3-264 of the 272-residue polypeptide. Electron density corresponding to the uncleaved 21-residue N-terminal histidine tag was not apparent for any subunit. The C-terminal tail adopts two conformations, with the tails of two monomers in the trimer extending away from the trimer and the remaining monomer exhibiting a folded tail conformation. The two monomers in the hexamer (one from each trimer) with folded C-terminal tails had poor electron density for residues 250–254 towards the C-terminal end. Data show the presence of L-malate (present in the crystallisation condition) forming meaningful active site interactions in the structure.

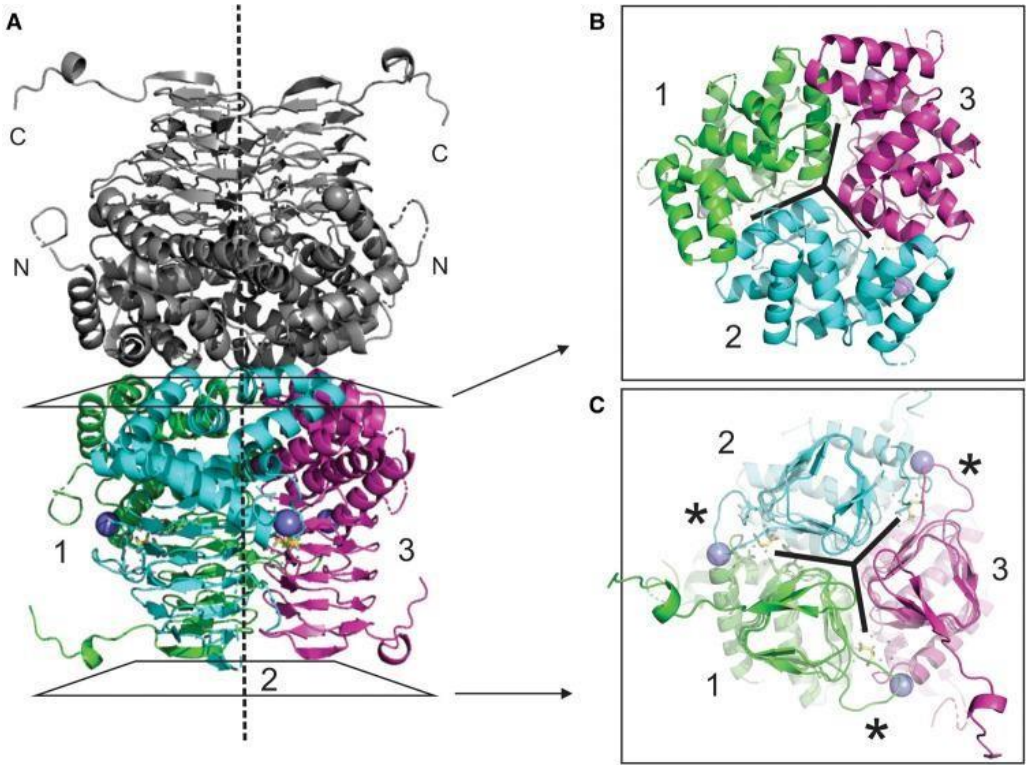


Figure 3.2 Structure of *Ng*CysE hexamer. A ribbon diagram of the overall structure of *Ng*CysE shows that a dimer of trimers forms the hexameric unit through interactions between the N-terminal α-helical domains (A). Trimers form through

interactions of the N-terminal α-helical domains (A). Trimers form through interactions of the N-terminal α-helices (B) and left-handed β-helical C-terminal domain (C). The vertical dashed line represents the independent three-fold symmetry which is evident in (B) and (C). Sodium ions are shown as a purple spheres and L-malate is shown as yellow sticks. Asterisks shows active site extended loop. This figure was prepared with PyMOL.

Table 3.2 Data collection and refinement statistics.

	***Ng*CysE**	***Ng*CysE + L-ser**
***Data collection*[a]**		
Wavelength (Å)	0.9537	0.9537
Resolution range (Å)	43.11 - 2.01 (2.08 - 2.01)	43.88-2.80 (2.8 - 2.95)
Space group	$P2_1$	C2
Unit cell parameters *a, b, c* (Å) *α , β, γ* (°)	77.90, 93.87, 102.02, 90.00, 91.29, 90.00	102.86, 94.59, 79.25, 90.00, 92.18, 90.00
No. of molecules in asymmetric unit	6	3
Total reflections	689150 (70393)	130367 (19576)
Unique reflections	97700 (9715)	18772 (2723)
Multiplicity	7.1 (7.2)	6.9 (7.2)
Completeness (%)	99.6 (100.0)	99.9 (100.0)
Mean I/σ(I)	18.5 (3.4)	9.1(2.2)
R_{merge}[b]	0.05574 (0.4889)	0.149 (0.915)
Refinement		
Reflections used in refinement	97327 (9715)	130367 (19576)
Reflections used for R_{free}	4749 (424)	18772 (2723)
R_{work}	0.1830 (0.2178)	0.2346 (0.2892)
R_{free}	0.2219 (0.2745)	0.2673 (0.3240)
No. protein atoms	11158	5001
No. solvent atoms	538	39
No. ligand atoms		
l-malate	54	-
sodium	6	-
l-serine	-	14
Protein residues	1534	726
r.m.s.d bonds (Å)	0.007	0.002
r.m.s.d angles (°)	0.84	0.59
Ramachandran favoured (%)	97.7	95.61
Ramachandran allowed (%)	2.3	4.39
Ramachandran outliers (%)	0	0

Average *B* (Å^2)	38.9	54.19
macromolecules	38.8	54.20
ligands	37.0	58.56
solvent	40.3	50.78
Clashscore	2.3	2.25
No. of TLS groups	1	4
PDB entry	6WYE	7RA4

[a]Statistics for highest resolution shell are shown in parentheses.

[b]$R_{merge} = \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hklj}$

The crystal structure of *Ng*CysE with serine bound (*Ng*CysE + L-Ser) was solved using a closed tail monomer of *Ng*CysE (6WYE) and was refined to 2.8 Å. The model has an R and free R of 0.234 and 0.267, respectively. A trimer is present in the asymmetric unit but a hexamer can be generated using the crystallographic twofold axis. There was density to support placement of substrate L-serine in two of the three active sites, with the third having weak density, where serine interacts with conserved active site residues.

There are no major conformational differences between the L-malate and L-serine bound structures (RMSD 0.236 Å), where the main differences are due to the presence of open tail conformations in *Ng*CysE. Density is present for extended and folded C-terminal tails in the *Ng*CysE + L-Ser structure (one closed and two open tails), but as a result of poorer resolution for this structure it was not possible to build the tails into the density. Given that the *Ng*CysE structure has better resolution, it was used for downstream structural analysis.

The CysE monomer can be subdivided into an N-terminal domain (residues 1–144) consisting of eight α-helices and a C-terminal left-handed β-helical domain (LβH; residues 145–272). Three monomers interact with each other through contacts between helices in the N-terminal domain and interactions of the LβH domain to form the trimer, which in turn form the hexamer as a dimer of trimers (Figure 3.2). The overall fold of the *Ng*CysE monomer and its hexameric state, is similar to the bacterial and plant homologues and strict residue conservation between species indicates the significance of this type of fold in CysEs ((Chattopadhyay et al., 2007), (Campanini et al., 2005; Mino et al., 1999), Figure 3.3).

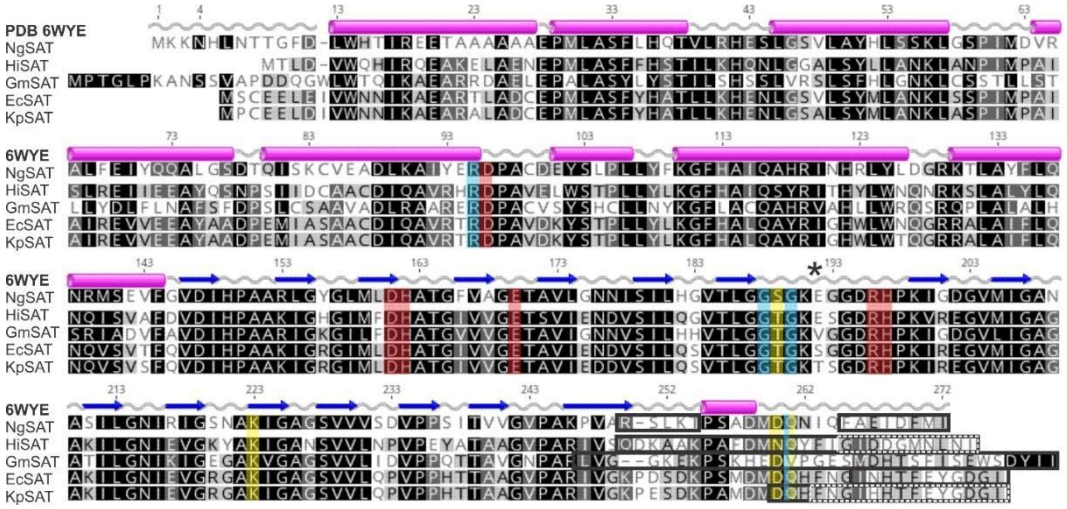


Figure 3.3 Sequence alignment of CysE enzymes. An amino acid sequence alignment of the most structurally similar CysEs to *N. gonorrhoeae* shows strict conservation of residues in the C-terminal LβH domain and in the active site. Secondary structure and numbering presented is from *N. gonorrhoeae* (6WYE) showing α-helices as pink cylinders and β-strands as blue arrows. Letter shading is based on % similarity. Red shading shows active site residues and blue shading shows additional important active site solvent interacting residues. Yellow shading shows residues identified in C-terminal tail folding and black (and dashed) boxes illustrate missing residues from these structures. *N. gonorrhoeae* (*Ng*CysE; AAW90067.1), *H. influenzae* (*Hi*CysE; WP_005694542.1), *G. max* (*Gm*CysE; XP_003528805.2), *E. coli* (*Ec*CysE; WP_114569552.1) and *K. pneumoniae* (*Kp*CysE; Q0ZB96). Asterisk shows the active site extended loop region. This figure was prepared with Geneious Prime 2019.2.1.

The acetyltransferases are characterised by a unique hexapeptide repeat ([LIV]-[GAED]-X^2-[STAV]-X) which results in the formation of the LβH domain. There is a single deviation from the LβH domain structure between β-strands 7 and 8, which creates an extended loop (residues 191–199 in this structure) which is significant as this forms one half of the active site (Figure 3.2C and 3.4). The active site is situated between the C-terminal domains of two adjacent monomers. In each trimer there are three active sites. In a trimer positioned with the N-terminus top and C-terminus bottom, the active site is located at the intersection of the N-terminal domain and C-terminal domain of the left monomer and the inner face of the extended loop from the right monomer (Figure 3.5). The six active site residues in *N. gonorrhoeae* comprise of Asp96, Asp161 and His162 from the left monomer and Asp147, Glu170, Arg196 and His197 from the right monomer (Figure 3.3 and Figure 3.4), with His162 and Asp147 forming the catalytic dyad.

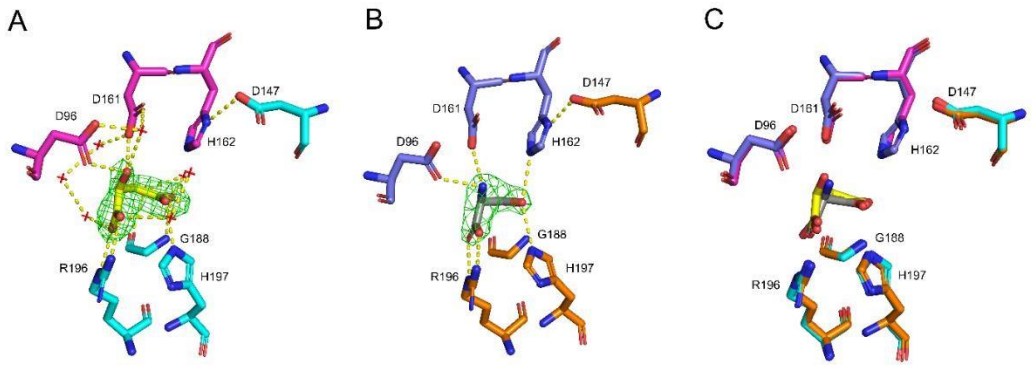


Figure 3.4. Simulated-annealing OMIT maps of ligand bound active sites for *Ng*CysE and *Ng*CysE + L-ser. The placement of L-malate (yellow sticks) in the *Ng*CysE active site (adjacent monomers coloured magenta and blue) (A) and L-serine (grey sticks) in the serine bound structure (adjacent monomers coloured purple and orange) (B) are supported by F_o-F_c maps displayed as green mesh (contoured at 3.0σ). (C) Overlay of *Ng*CysE and *Ng*CysE + L-ser monomers (open conformation) demonstrates conserved orientation of active site residues Asp96, Asp147, Asp161, His162, Gly188, Arg196 and His197 (displayed as sticks). Hydrogen bonds between active site residues and ligands are represented by yellow dashed lines. Waters are shown as red crosses in (A) no waters are shown in (B) due to the difficult placement of waters due to resolution. This figure was prepared using PyMOL.

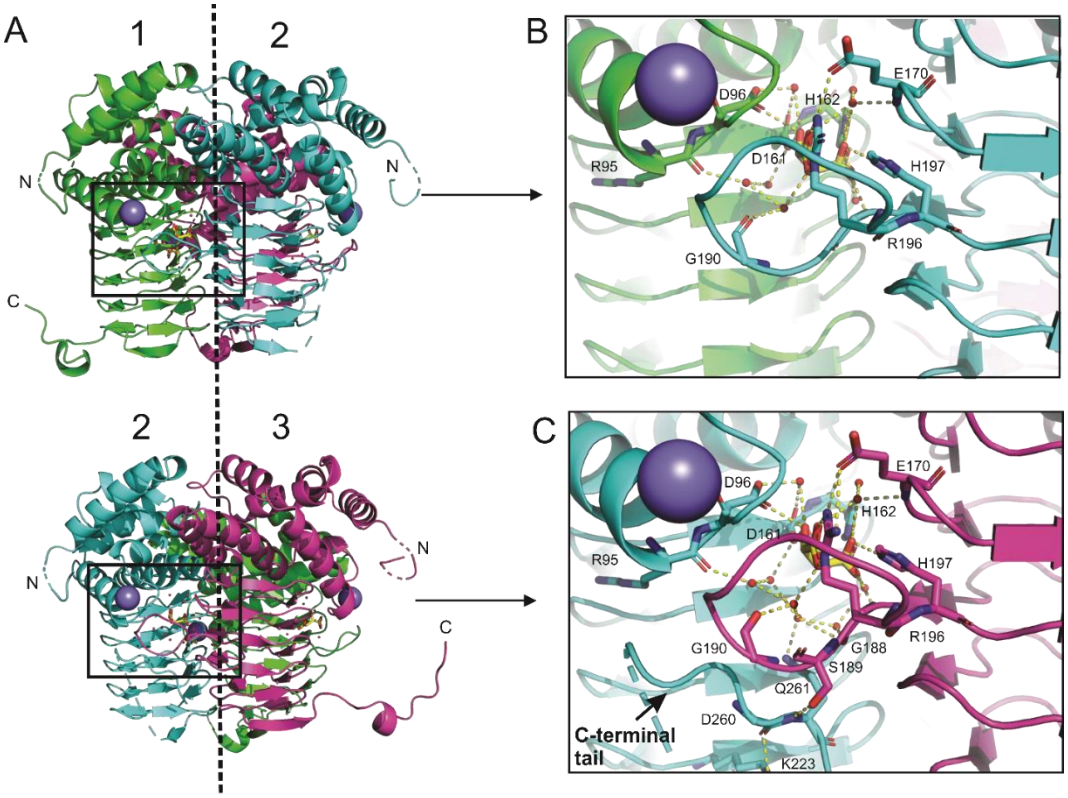


Figure 3.5 A ribbon diagram of the *Ng*CysE trimer illustrates the position and formation of the active sites between two adjacent monomers. The structure reveals different C-terminal tail conformations in different active sites of a single trimer (A). Between monomers 1 and 2 the C-terminal tail of monomer 1 (green)

is extended away from the active site (B). Between monomers 2 and 3 the C-terminal tail of monomer 2 (blue) is tucked up into the active site (C). Sodium ions are shown as a purple spheres and L-malate is shown as yellow sticks. Dashed line represents the 3-fold symmetry. This figure was prepared with PyMOL.

Comparing the two *Ng*CysE structures, the active site residues, Asp96, Asp161, Arg196 and His197 (Figure 3.4), interact with L-malate in a similar manner as the substrate L-serine. The terminus carboxylic acid group on L-malate and L-serine are co-ordinated by the Arg196 side chain. The negatively charged carboxylate side chains from residues Asp96 and Asp161 both co-ordinate the L-serine backbone amide and the hydroxyl group in L-malate (Figure 3.4A). The most notable difference in active site coordination is where the L-serine γ hydroxyl forms meaningful hydrogen bonds to flanking His162 and His197 (Figure 3.4B), whereas in the *Ng*CysE malate bound structure the equivalent carboxylic acid group maintains a hydrogen bond with His197 but is orientated away from the catalytic His162, and is co-ordinated by the backbone amide of Gly188 (Figure 3.4). This difference is likely due to L-serine being a natural substrate of the enzyme, whereas L-malate is not. The interaction of L-serine with active site residues in the *Ng*CysE + L-Ser structure is similar to the other L-serine bound CysE structure in the PDB (4N69) from Soybean (Glycine max) (Yi et al., 2013), with the exception that His162 is interacting with L-serine in the *Ng*CysE + L-Ser structure.

The C-terminal tail adopts two conformations in our malate bound structure. For two monomers, the tails extends outward away from the trimer, reaching towards adjacent trimers in the crystal unit and occupying the interface between two monomers (exhibiting domain swapping). For the remaining single monomer in each trimer, the tail loops upward against the side of the monomer and tucks under the active site (Figure 3.5). There are notable differences in the number of water molecules participating in the active site chemistry with L-malate when the tail is positioned near the active site (Figure 3.5B and C). In the extended tail conformation with an ‘open’ active site there are seven water molecules interacting with L-malate, whereas in the tucked tail conformation with a ‘closed’ active site there are an additional two water molecules coordinating with L-malate (Figure 3.5B and C). The ‘closed’ active site recruits additional residues Gly188 and Gln261 to join Arg95 and Gly190 in solvent interactions and residues Ser189, Lys223, Asp260 and Gln261 become involved in tail interactions (Figure 3.5C).

Glutamine 261 is noted to be of particular importance in this 'closed' active site conformation as this has key interactions securing the tail position and involvement in the active site chemistry, drawing in additional water molecules. Notably the L-malate bound structure presented here has only eight missing C-terminal residues compared with related structures (Figure 3.3) that have more missing residues. When the active site is in the 'closed' conformation there is no access to the interior serine binding pocket.

The tucked tail conformation of the *Ng*CysE monomer is identical with homologous CysE structures that have the product cysteine bound (PDB 1SSQ, 1T3D, 3GVD, 3P47, 4H7O and 6JVU) and analysis of these structures also reveals that there would be no access to the pocket of the active site (regardless of the number of missing C-terminal residues). Due to an artifact of crystallographic packing, we have surprisingly been able to examine two active site states regardless of having a true substrate molecule bound. L-malate was present in the crystallisation condition and its occupation of the active site is expected to be based on charge in a similar manner to the inhibitor L-cysteine. Comparing the 'closed' active site of *Ng*CysE with the homologous *E. coli* CysE (1T3D) (containing cysteine in the active site) we observe a similar arrangement of active site residues and the C-terminal tail position. Cysteine acts as a negative feedback regulator of CysE activity and it could be deduced that it does this by 'closing' the active site. Comparing the 'open' active site of *Ng*CysE with the *H. influenzae* homologue (1SST) (containing CoA in the active site) we observe that the C-terminal tail is also orientated away from the active site and in addition the extended loop in 1SST is disrupted whereas in our structure is not. The substrate CoA is a large molecule and by occupying the active site the C-terminal tail cannot fold into this space. The structure presented here is unique in that we see both 'open' and 'closed' conformations in a monomer which gives insight into substrate binding.

3.5.4 Cysteine inhibition

For CysE enzymes, feedback inhibition by L-cysteine is well-documented across both bacterial and plant species (Chronis & Krishnan, 2004; Johnson et al., 2004b; Kredich & Tomkins, 1966). Here, we characterised L-cysteine inhibition of *Ng*CysE, relative to the substrates L-serine and acetyl-CoA. *Ng*CysE activity was

inhibited by increasing concentrations of L-cysteine with an IC_{50} of 6.4 ± 0.1 µM and a hill slope of −1.86 ± 0.15 (Figure 3.4A). Collection of Michaelis–Menten plots with varying L-serine concentrations (Figure 3.4B) in the presence of fixed concentrations of cysteine, increased the K_M (serine) from 1.16 mM with no L-cysteine bound to 4.83 mM in the presence of 8 µM L-cysteine, while the k_{cat} did not change. This demonstrates that L-cysteine is a competitive inhibitor relative to L-serine. Competitive inhibition is supported by structural evidence from *E. coli* and *E. histolytica* where cysteine has been shown to bind directly to the L-serine active site, preventing L-serine from binding (Kumar et al., 2011; Pye et al., 2004). We could not confidently fit competitive, non-competitive or uncompetitive enzyme inhibition models to the data. Incorporation of the hill coefficient (Equation A.1) (Weiss, 1977; Aragao et al., 2018) into the competitive inhibition model accounts for cooperativity of binding and is well accounted for in the data (Figure A.5). The h value (a measure of cooperativity) decreases steadily from 1 (noncooperative binding) to 0.299 (negatively cooperative binding) with increasing L-cysteine concentrations (Table A.1), suggesting negative cooperativity between active sites when L-cysteine is present. This is consistent with the hill slope (−1.86) of the L-cysteine dose response curve (Figure 3.6A), where binding of L-cysteine increases the affinity of other active sites in the multimer for L-cysteine (positive cooperativity) which reduces the affinity for L-serine (negative cooperativity in Figure 3.6B and Figure A.5). Although cooperativity has not been reported for CysE homologues, together with the positive cooperativity seen in the CoA dose response curve (Figure 3.1C), it clear that there is interconnectedness between active sites in *Ng*CysE, which is unsurprising given the multimeric nature of the enzyme.

L-cysteine has also been shown to be a competitive inhibitor relative to acetyl-CoA in CysE enzymes from other bacteria (Kumar et al., 2011; Pye et al., 2004). To investigate if L-cysteine is a competitive inhibitor relative to acetyl-CoA in *Ng*CysE we collected Michaelis-Menten plots with varying acetyl-CoA concentrations (Figure 3.6C) in the presence of 4, 6 and 8 µM L-cysteine respectively. We were unable to confidently determine the mode of inhibition, with both competitive and non-competitive inhibition models fitting the data reasonably. Incorporation of the hill coefficient into either equation did not improve the fit of either model.

Collection of IC_{50} curves in the presence of increasing concentrations of acetyl-CoA (Figure 3.6D), demonstrate an increase in IC_{50} from 6.3 µM (hill coefficient 2.1 ± 0.2) to 10.8 µM (hill coefficient 2.5 ± 0.1), consistent with a competitive inhibition mechanism (Rosa et al., 2019) and positive cooperative for L-cysteine binding. Competitive inhibition is supported by the observation that in cysteine-bound CysE structures from other bacteria, the flexible C-terminal tail of CysE is folded against the active site (Kumar et al., 2011; Olsen et al., 2004; Pye et al., 2004), as also seen in the 'closed' active sites in our *Ng*CysE structure. This essentially closes the active site, preventing acetyl-CoA from binding, thereby acting as a competitive inhibitor.

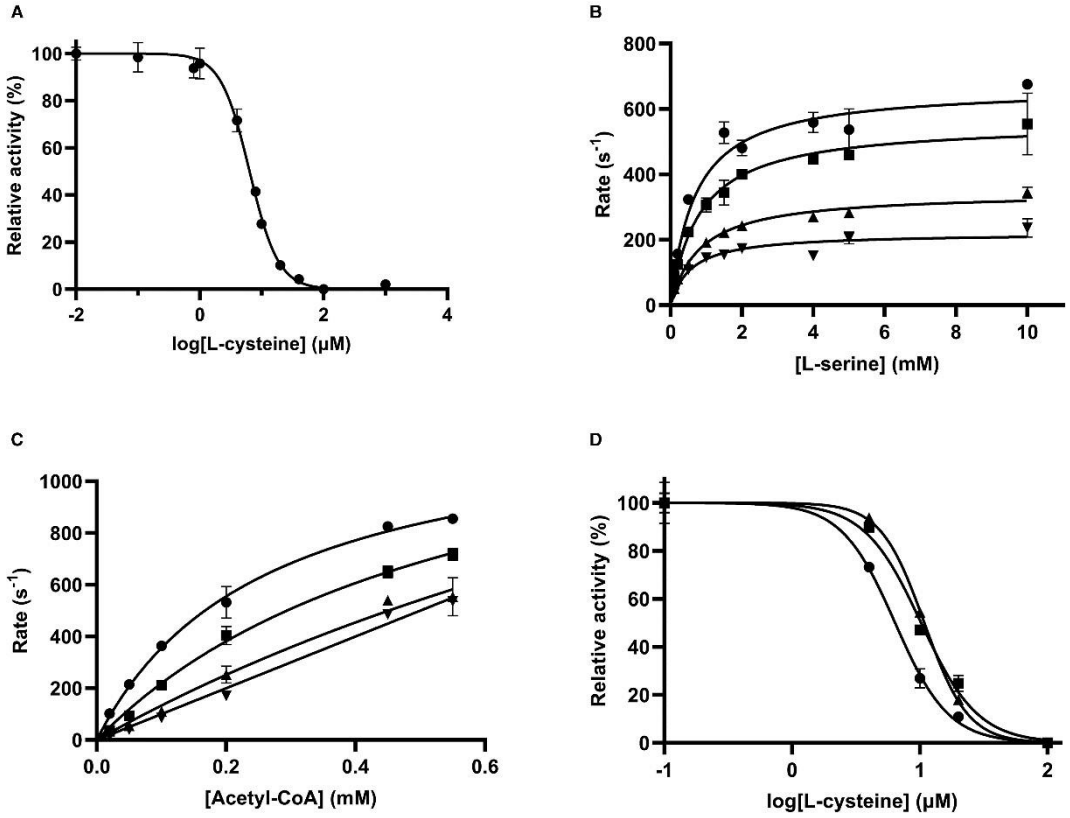


Figure 3.6 Characterisation of cysteine inhibition. (A) Dose response curve for L-cysteine. Line represents the fit of IC_{50} equation to data. Substrate concentrations were fixed at 0.15 mM and 1.5 mM, for acetyl-CoA and L-serine, respectively. (B) Michaelis-Menten plots were collected for 0 (circles), 4 (squares), 6 (triangles) and 8 (inverted triangles) µM of L-cysteine, with varying concentrations of serine. (C) Michaelis-Menten plots for 0 (circles), 4 (squares), 6 (triangles) and 8 (inverted triangles) µM of L-cysteine, with varying concentrations of acetyl-CoA (x-axis). (D) Dose-response curves for L-cysteine in the presence of 0.15 (circles), 0.25 (squares) and 0.45 (triangles) mM of acetyl-CoA. Data points are mean and error bars are SEM derived from two replicates.

3.5.5 Small angle X-ray scattering shows a conformational change in the presence of L-cysteine

SEC in line with small angle X-ray scattering (SEC-SAXS) analysis (Kirby et al., 2016) was performed to determine the effect of L-cysteine binding to *Ng*CysE in solution. Crystallisation attempts of *Ng*CysE with L-cysteine were unsuccessful due to degradation of crystals upon addition of L-cysteine. The scattering data were consistent with a globular protein (Figure 3.5), and the calculated molecular mass based on the Porod volumes, combined with the pair wise distribution (P(r)) analysis support a hexamer of *Ng*CysE monomers. The calculated molecular mass for the L-cysteine bound and unbound structures was the same (193.87 kDa). The molecular mass of the monomer including the His-tag is 31.6 kDa, therefore, 193.87 kDa divided by the molecular mass of the monomer (31.6 kDa) is 6.1 monomers per structure, consistent with the hexameric crystal structure and SEC presented above. The SAXS profiles were fit with the calculated theoretical scattering from our *Ng*CysE crystal structure. The theoretical scattering profile shows a poor fit with the scattering data collected for unbound (χ^2 = 5.31) and L- cysteine bound *Ng*CysE scattering data (χ^2 = 10.84) (Figure 3.5A), which is not unexpected due to combination of closed (tucked C-terminal tail) and open (flexible C-terminal tail) conformations seen in the *Ng*CysE crystal structure which the theoretical scattering is calculated from.

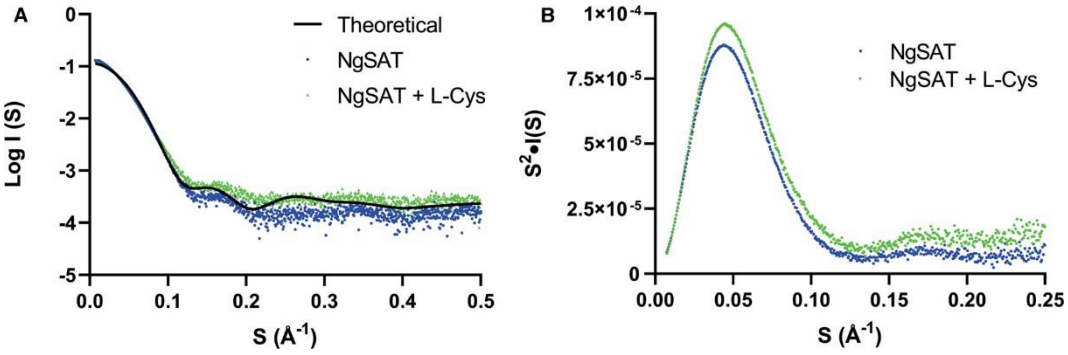


Figure 3.7 SAXS profiles of *Ng*CysE bound and unbound to L-cysteine. (A) SAXS profiles of *Ng*CysE were measured in the absence of (blue circles) and presence (green circles) of 50 µM L-cysteine (L-Cys). Theoretical scattering (black line) calculated from cysteine-free *Ng*CysE crystal structure. (B) Kratky plot of *Ng*CysE in the absence (blue circles) and presence (green circles) of 50 µM L-cysteine.

Comparison of the SAXS profiles in the presence and absence of L-cysteine shows clear differences (Figure 3.7A), consistent with a conformational change upon

binding of L-cysteine. The difference is associated with an increased compactness of the enzyme when bound to L-cysteine as shown by the smaller radius of gyration derived from the Guinier Plot (R*g* = 40.26 ± 0.35 Å with L-cysteine bound compared with R*g* = 43.01 ± 0.27 Å without) and P(r) analysis (R*g* = 40.8 ± 0.14 Å for L-cysteine bound compared with R*g* = 43.17 ± 0.14 Å for unbound). The scattering data, including the pairwise distribution analysis and R*g* values from the Guinier analysis are consistent with SAXS data from CysE from *E. coli* (Mino, Hiraoka, et al., 2000).

The bell-shaped curve observed in the Kratky plot (Figure 3.5B) shows that the protein exhibits a typical folded shape. This confirms the protein remains folded in the presence of L-cysteine although the cysteine-free curve is slightly flattened suggesting slightly increased flexibility of the protein. Decreased flexibility and an increased compactness of the cysteine bound enzyme is consistent with previously reported CysE structures crystallised with the cysteine inhibitor (Kumar et al., 2011; Olsen et al., 2004; Pye et al., 2004). The flexible C-terminal tail of CysE in cysteine bound structures adopts a more rigid, less flexible structure by interacting with the third β-coil to bury the acetyl-CoA binding site (Kumar et al., 2011) which is also consistent with the 'closed' state for *Ng*CysE that we see in our crystal structure leading to small conformational changes and a slightly more compact enzyme.

3.6 Conclusions

Cysteine synthesis in bacteria is a two-step reaction, the first step catalysed by CysE that produces *O*-acetylserine from L-serine and acetyl-CoA, and the second step catalysed by *O*-acetylserine sulphydrylase (OASS) that uses the substrates sulphide or thiosulphate to synthesise L-cysteine from *O*-acetylserine. The CysE and CysK enzymes form the cysteine synthase complex with the C-terminal tail of CysE binding into and inhibiting the active site of OASS-A. CysE is feedback inhibited by the end product of the pathway L-cysteine. Bacterial members of the CysE family are generally hexamers (with the exception of *E. histolytica* (Kumar et al., 2011)) and are characterised by a unique left-handed parallel β-helix (LβH).

Here we have presented structural and kinetic data for CysE from *N. gonorrhoeae* (*Ng*CysE). SEC and small angle X-ray scattering (SAXS) analysis confirm that *Ng*CysE is a hexamer in solution as well as in the crystal structure, with no other

oligomeric states detected. We have characterised the kinetics of *Ng*CysE, and hypothesise a sequential mechanism, although the order of substrate binding is unknown. We suggest substrate inhibition at concentrations of acetyl-CoA ≥ 1 mM but due to limitations in experimental set-up were not able to confirm substrate inhibition. This is in contrast with other CysE enzymes from other bacteria including *E. coli* and *H. influenzae* that display substrate inhibition by L-serine (Olsen et al., 2004; Pye et al., 2004). *Ng*CysE also exhibits product inhibition by co-enzyme A (CoA) consistent with other CysE enzymes (Hindson & Shaw, 2003) (Johnson et al., 2004b). Cooperativity is seen in the presence of the inhibitor L-cysteine and the product CoA indicating an interconnectedness between active sites in the *Ng*CysE hexamer. *Ng*CysE is feedback inhibited by the end product of the pathway L-cysteine, and using SAXS we show small conformational changes in the cysteine bound enzyme. These changes suggest a less flexible, slightly more compact enzyme when bound to cysteine. This is consistent with the 'closed' active sites we see in the *Ng*CysE crystal structure and crystal structures of other CysE enzymes bound to L-cysteine (Kumar et al., 2011; Olsen et al., 2004; Pye et al., 2004), where the C-terminal tail is tucked up against the side of the active site in the closed (cysteine-bound) structures leading to a more rigid enzyme.

The structure of *Ng*CysE reveals that this enzyme is clearly part of the CysE family due to its hexameric nature and LβH. As with other related structures, *Ng*CysE is a dimer of trimers, with each monomer comprising an N-terminal α-helical domain and a C-terminal LβH domain. The majority of CysE three-dimensional structures are incomplete and have up to 37 amino acids missing from the C-terminal end of the enzyme. We note that the *Ng*CysE structure has the least number of missing residues from the distal C-terminal end compared with related structures, which has allowed us to investigate the topology of the active site with some confidence. A crystallographic artifact has led to two different C-terminal tail conformations in the *Ng*CysE structure. Where other structures have missing residues, leading to an exposed active site cavity, *Ng*CysE exhibits either a free C-terminal tail swept outwards from the enzyme or a constrained C-terminal tail tucked up into the active site. The outward C-terminal tail results in an 'open' active site, in contrast with the tucked tail which results in a 'closed' active site. The ligand L-malate is present in the active site of *Ng*CysE due to its presence in the crystallography reagent. It

occupies the same area and interacts with the active site residues that the natural substrate L-serine would, as seen in our *Ng*CysE + Ser structure, with the exception of catalytic His162.

The tucked C-terminal tail 'closed' conformation reveals interesting interactions with both the solvent and other residues. A key glutamine residue (Gln261) in the tail, interacts with two additional water molecules in the active site (assisted by a nearby Gly188) that were otherwise uninvolved in the 'open' conformation. The interaction of Gln261 in the C-terminal tail brings a serine residue near the active site into close-range to allow it to form polar contacts with an aspartate on the C-terminal tail, which in turn forms a polar contact with a lysine from a β-strand in the LβH. Analysis of the residues involved in these polar contacts to tuck the C-terminal tail into the active site reveals strict conservation between species. Overall, there is a chain reaction of polar contacts from the active site, downwards through the tail, securing it and closing the active site. In bacteria the primary pathways of inorganic sulphur assimilation converge at cysteine biosynthesis and *N. gonorrhoeae* lacks the ability to utilise sulphate as an inorganic sulphur resource, but its requirement for sulphur can be fulfilled by thiosulphate (Le Faou, 1984). Curiously *N. gonorrhoeae* lacks the variant of the OASS enzyme that utilises thiosulphate (OASS-B/CysM), having just a single OASS enzyme with homology to OASS-A/CysK (the variant that utilises sulphide not thiosulphate) (Hicks & Mullholland, 2018). It remains to be seen if this enzyme does indeed utilise only sulphide as a substrate or if it is a dual-function enzyme, also capable of using thiosulphate for the synthesis of cysteine. In addition it is unknown if OASS-A from *N. gonorrhoeae* interacts with *Ng*CysE to form the cysteine synthase complex although it is likely, as the C-terminal tail of *Ng*CysE contains the conserved terminal isoleucine critical for binding of the tail into the active site of OASS (Campanini et al., 2005; Mino, Hiraoka, et al., 2000; Mino et al., 1999).

Our *Ng*CysE structures presented here, alongside kinetic and inhibition data, gives unique insight into the structure, function and inhibition of CysE from *N. gonorrhoeae* and due to the unique nature of sulphur acquisition for cysteine biosynthesis in *N. gonorrhoeae* could represent a novel drug target for treating extensively antimicrobial-resistant gonorrhoea.

3.7 Supplementary Materials

Supplementary materials associated with this chapter are available in Appendix A: Supplementary material for Chapter Three.

3.8 Data availability

The final co-ordinate and structure amplitude files for *Ng*CysE and *Ng*CysE + L-Ser are deposited in the Protein Data Bank (https://www.rcsb.org/) under PDB accession codes 6WYE and 7RA4, respectively. UniProt Accession ID for NGFG_01496 (*Ng*CysE): Q5F6X0 (Q5F6X0_NEIG1).

3.9 References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., . . . Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 2), 213-221. 10.1107/S0907444909052925

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., . . . Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography*, *68*(Pt 4), 352-367. 10.1107/S0907444912001308

Agarwal, S. M., Jain, R., Bhattacharya, A., & Azam, A. (2008). Inhibitors of Escherichia coli serine acetyltransferase block proliferation of Entamoeba histolytica trophozoites. *International Journal for Parasitology, 38*(2), 137-141. 10.1016/j.ijpara.2007.09.009

Aragao, D., Aishima, J., Cherukuvada, H., Clarken, R., Clift, M., Cowieson, N. P., . . . Caradoc-Davies, T. T. (2018). MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *Journal of Synchrotron Radiation, 25*(3), 885-891. doi:10.1107/S1600577518003120

Benoni, R., De Bei, O., Paredi, G., Hayes, C. S., Franko, N., Mozzarelli, A., . . . Campanini, B. (2017). Modulation of Escherichia coli serine acetyltransferase catalytic activity in the cysteine synthase complex. *FEBS letters, 591*(9), 1212-1224. 10.1002/18733468.12630

Berney, M., Berney-Meyer, L., Wong, K. W., Chen, B., Chen, M., Kim, J., . . . Jacobs, W. R., Jr. (2015). Essential roles of methionine and S-adenosylmethionine in the autarkic lifestyle of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A, 112*(32), 10008-10013. 10.1073/pnas.1513033112

Campanini, B., Pieroni, M., Raboni, S., Bettati, S., Benoni, R., Pecchini, C., . . . Mozzarelli, A. (2015). Inhibitors of the Sulphur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Current Medicinal Chemistry, 22*(2), 187-213. 10.2174/0929867321666141112122553

Campanini, B., Speroni, F., Salsi, E., Cook, P. F., Roderick, S. L., Huang, B., . . . Mozzarelli, A. (2005). Interaction of serine acetyltransferase with O-acetylserine sulfhydrylase active site: evidence from fluorescence spectroscopy. *Protein science : a publication of the Protein Society*, *14*(8), 2115-2124. 10.1110/ps.051492805

Chattopadhyay, A., Meier, M., Ivaninskii, S., Burkhard, P., Speroni, F., Campanini, B., . . . Cook, P. F. (2007). Structure, Mechanism, and Conformational Dynamics of

OAcetylserine Sulfhydrylase from Salmonella typhimurium: Comparison of A and B Isozymes. *Biochemistry, 46*(28), 8315-8330. 10.1021/bi602603c

Chen, C., Han, X., Yan, Q., Wang, C., Jia, L., Taj, A., . . . Ma, Y. (2019). The Inhibitory Effect of GlmU Acetyltransferase Inhibitor TPSA on Mycobacterium tuberculosis May Be Affected Due to Its Methylation by Methyltransferase Rv0560c. *Frontiers in Cellular and Infection Microbiology, 9* 10.3389/fcimb.2019.00251

Chronis, D., & Krishnan, H. B. (2004). Sulphur assimilation in soybean (Glycine max [L.] Merr.): molecular cloning and characterization of a cytosolic isoform of serine acetyltransferase. *Planta, 218*(3), 417-426. 10.1007/s00425-003-1123-3

Cook, P. F., & Wedding, R. T. (1977). Initial kinetic characterization of the multienzyme complex, cysteine synthetase. *Archives of Biochemistry and Biophysics, 178*(1), 293-302. 10.1016/0003-9861(77)90194-1

Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D, 60*(12 Part 1), 2126-2132. doi:10.1107/S0907444904019158

Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D, 69*(7), 1204-1214. doi:10.1107/S0907444913000061

Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., . . . Svergun, D. I. (2017). ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr, 50*(Pt 4), 1212-1225. 10.1107/s1600576717007786

Gesztelyi, R., Zsuga, J., Kemeny-Beke, A., Varga, B., Juhasz, B., & Tosaki, A. (2012). The Hill equation and the origin of quantitative pharmacology. *Archive for History of Exact Sciences, 66*(4), 427-438. 10.1007/s00407-012-0098-5

Hicks, J. L., & Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology, 164*(12), 1471-1480. 10.1099/mic.0.000728

Hindson, V. J. (2003). Serine acetyltransferase of Escherichia coli: substrate specificity and feedback control by cysteine. *The Biochemical journal, 375*(Pt 3), 745-752. 10.1042/BJ20030429

Hindson, V. J., & Shaw, W. V. (2003). Random-Order Ternary Complex Reaction Mechanism of Serine Acetyltransferase from Escherichia coli. *Biochemistry, 42*(10), 3113-3119. 10.1021/bi0267893

Hirai, M. Y., Fujiwara, T., Awazuhara, M., Kimura, T., Noji, M., & Saito, K. (2003). Global expression profiling of sulphur-starved Arabidopsis by DNA macroarray reveals the role of O-acetyl-l-serine as a general regulator of gene expression in response to sulphur nutrition. *Plant J, 33*(4), 651-663. 10.1046/j.1365313x.2003.01658.x

Johnson, C. M., Huang, B., Roderick, S. L., & Cook, P. F. (2004a). Chemical Mechanism of the Serine Acetyltransferase from Haemophilus influenzae. *Biochemistry, 43*(49), 15534-15539. 10.1021/bi048450h

Johnson, C. M., Huang, B., Roderick, S. L., & Cook, P. F. (2004b). Kinetic mechanism of the serine acetyltransferase from Haemophilus influenzae. *Archives of Biochemistry and Biophysics, 429*(2), 115-122. 10.1016/j.abb.2004.06.006

Johnson, C. M., Roderick, S. L., & Cook, P. F. (2005). The serine acetyltransferase reaction: acetyl transfer from an acylpantothenyl donor to an alcohol. *Archives of Biochemistry and Biophysics, 433*(1), 85-95. 10.1016/j.abb.2004.08.014

Kabsch, W. (2010). XDS. *Acta crystallographica. Section D, Biological crystallography, 66*(Pt 2), 125-132. 10.1107/S0907444909047337

Kirby, N., Cowieson, N., Hawley, A. M., Mudie, S. T., McGillivray, D. J., Kusel, M., . . . Ryan, T. M. (2016). Improved radiation dose efficiency in solution SAXS using a sheath flow sample environment. *Acta Crystallogr D Struct Biol, 72*(Pt 12), 12541266. 10.1107/s2059798316017174

Kredich, N. M., & Tomkins, G. M. (1966). The Enzymic Synthesis of l-Cysteine in Escherichia coli and Salmonella typhimurium. *Journal of Biological Chemistry, 241*(21), 4955-4965.

Kumar, S., Raj, I., Nagpal, I., Subbarao, N., & Gourinath, S. (2011). Structural and Biochemical Studies of Serine Acetyltransferase Reveal Why the Parasite Entamoeba histolytica Cannot Form a Cysteine Synthase Complex. *Journal of Biological Chemistry, 286*(14), 12533-12541. 10.1074/jbc.M110.197376

Kumaran, S., Yi, H., Krishnan, H. B., & Jez, J. M. (2009). Assembly of the cysteine synthase complex and the regulatory role of protein-protein interactions. *J Biol Chem, 284*(15), 10268-10275. 10.1074/jbc.M900154200

Le Faou, A. (1984). Sulphur nutrition and metabolism in various species of Neisseria. *Annales de l'Institut Pasteur / Microbiologie, 135*(1, Supplement B), 3-11. 10.1016/S0769-2609(84)80037-X

Leu, L. S., & Cook, P. F. (1994). Kinetic Mechanism of Serine Transacetylase from Salmonella typhimurium. *Biochemistry, 33*(9), 2667-2671. 10.1021/bi00175a040 Levticus 15:1-3. *Holy Bible* (King James Version ed.).

Magalhães, J., Franko, N., Raboni, S., Annunziato, G., Tammela, P., Bruno, A., . . . Costantino, G. (2020). Inhibition of Nonessential Bacterial Targets: Discovery of a Novel Serine O-Acetyltransferase Inhibitor. *ACS Medicinal Chemistry Letters* 10.1021/acsmedchemlett.9b00627

Matthews, B. W. (1968). Solvent Content of Protein Crystals. *Journal of Molecular Biology, 33*(2), 491-497. 10.1016/0022-2836(68)90205-2

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography, 40*(Pt 4), 658-674. 10.1107/S0021889807021206

Mino, K., Hiraoka, K., Imamura, K., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2000). Characteristics of serine acetyltransferase from Escherichia coli deleting different lengths of amino acid residues from the C-terminus. *Biosci Biotechnol Biochem, 64*(9), 1874-1880. 10.1271/bbb.64.1874

Mino, K., Yamanoue, T., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (1999). Purification and Characterization of Serine Acetyltransferase from Escherichia coli Partially Truncated at the C-Terminal Region. *Bioscience, Biotechnology, and Biochemistry, 63*(1), 168-179. 10.1271/bbb.63.168

Murima, P., McKinney, John D., & Pethe, K. (2014). Targeting Bacterial Central Metabolism for Drug Development. *Chemistry & Biology, 21*(11), 1423-1432. 10.1016/j.chembiol.2014.08.020

Noji, M., Inoue, K., Kimura, N., Gouda, A., & Saito, K. (1998). Isoform-dependent Differences in Feedback Regulation and Subcellular Localization of Serine Acetyltransferase Involved in Cysteine Biosynthesis from Arabidopsis thaliana. *Journal of Biological Chemistry, 273*(49), 32739-32745.10.1074/jbc.273.49.32739

Olsen, L., B., H., Vetting, M., & Roderick, S. (2004). Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor. *Biochemistry, 43*(20), 6013-6019.

Olsen, L. R., Vetting, M. W., & Roderick, S. L. (2007). Structure of the E. coli bifunctional GlmU acetyltransferase active site with substrates and products. *Protein science : a publication of the Protein Society, 16*(6), 1230-1235. 10.1110/ps.072779707

Pye, V. E., Tingey, A. P., Robson, R. L., & Moody, P. C. E. (2004). The Structure and Mechanism of Serine Acetyltransferase from Escherichia coli. *Journal of Biological Chemistry, 279*(39), 40729-40736. 10.1074/jbc.M403751200

Rabeh, W. M., & Cook, P. F. (2004). Structure and Mechanism of O-Acetylserine Sulfhydrylase. *Journal of Biological Chemistry, 279*(26), 26803-26806. 10.1074/jbc.R400001200

Raetz, C. R. H., & Roderick, S. L. (1995). A Left-Handed Parallel β Helix in the Structure of UDP-N-Acetylglucosamine Acyltransferase. *Science, 270*(5238), 997-1000. 10.1126/science.270.5238.997

Remmele, C. W., Xian, Y., Albrecht, M., Faulstich, M., Fraunholz, M., Heinrichs, E., . . . Rudel, T. (2014). Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic acids research, 42*(16), 10579-10595. 10.1093/nar/gku762

Rosa, B., Marchetti, M., Paredi, G., Amenitsch, H., Franko, N., Benoni, R., . . . Bettati, S. (2019). Combination of SAXS and Protein Painting Discloses the ThreeDimensional Organization of the Bacterial Cysteine Synthase Complex, a Potential Target for Enhancers of Antibiotic Action. *International Journal of Molecular Sciences, 20*(20), 5219.

Ryan, T. M., Trewhella, J., Murphy, J. M., Keown, J. R., Casey, L., Pearce, F. G., . . . Kirby, N. (2018). An optimized SEC-SAXS system enabling high X-ray dose for rapid SAXS assessment with correlated UV measurements for biomolecular structure analysis. *Journal of Applied Crystallography, 51*(1), 97-111. 10.1107/S1600576717017101

Schurr, J. M., & Schmitz, K. S. (1976). Orientation constraints and rotational diffusion in bimolecular solution kinetics. A simplification. *The Journal of Physical Chemistry, 80*(17), 1934-1936. 10.1021/j100558a026

Svergun, D., Barberato, C., & Koch, M. H. J. (1995). CRYSOL– a Program to Evaluate Xray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography, 28*(6), 768-773. 10.1107/S0021889895007047

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., . . . Adams, P. D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta crystallographica. Section D, Biological crystallography, 64*(Pt 1), 61-69. 10.1107/S090744490705024X

Terwilliger, T. C., Klei, H., Adams, P. D., Moriarty, N. W., & Cohn, J. D. (2006). Automated ligand fitting by core-fragment fitting and extension into density. *Acta crystallographica. Section D, Biological crystallography, 62*(Pt 8), 915-922. 10.1107/S0907444906017161

Venkatachalam, C. M., Jiang, X., Oldfield, T., & Waldman, M. (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling, 21*(4), 289-307. 10.1016/S10933263(02)00164-X

Verma, D., Gupta, S., Saxena, R., Kaur, P., R, R., Srivastava, S., & Gupta, V. (2020). Allosteric inhibition and kinetic characterization of Klebsiella pneumoniae CysE: An emerging drug target. *Int J Biol Macromol, 151*, 1240-1249. 10.1016/j.ijbiomac.2019.10.170

Weiss, J. N. (1997). The Hill equation revisited: uses and misuses. *The FASEB Journal, 11*(11), 835-841. 10.1096/fasebj.11.11.9285481

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., . . . Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography, 67*(Pt 4), 235-242. 10.1107/S0907444910045749

Wirtz, M., & Hell, R. (2006). Functional analysis of the cysteine synthase protein complex from plants: Structural, biochemical and regulatory properties. *Journal of Plant Physiology, 163*(3), 273-286. 10.1016/j.jplph.2005.11.013

Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B., & Jez, J. M. (2013). Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. *The Journal of biological chemistry, 288*(51), 36463-36472. 10.1074/jbc.M113.527143

Zhao, C., Kumada, Y., Imanaka, H., Imamura, K., & Nakanishi, K. (2006). Cloning, overexpression, purification, and characterization of O-acetylserine sulfhydrylase-B from Escherichia coli. *Protein Expr Purif*,*47*(2), 607-613. 10.1016/j.pep.2006.01.002

Zheng, H., Cooper, D. R., Porebski, P. J., Shabalin, I. G., Handing, K. B., & Minor, W. (2017). CheckMyMetal: a macromolecular metal-binding validation tool. *Acta Crystallographica Section D, 73*(3), 223-233. doi:10.1107/S2059798317001061

Chapter Four: Identification of Inhibitors of Serine Acetyltransferase from *Neisseria gonorrhoeae*

4.1 Preface

Targeting *de novo* cysteine biosynthesis for antimicrobial development has garnered interest, particularly for combatting antimicrobial-resistant pathogens (Campanini et al., 2015; Chen et al., 2019; Gupta & Gupta, 2021; Toyomoto et al., 2023; Verma et al., 2020). This chapter builds upon previous work in Chapter Three where we presented the kinetic and structural characterisation of CysE from *Neisseria gonorrhoeae* (*Ng*CysE) (Oldham, Prentice, Summers, & Hicks, 2022). In this chapter, we have used the crystal structure of *Ng*CysE (PDB 6WYE) to conduct structure-based virtual inhibitor screening from commercial libraries, ordered hit compounds identified *in silico* and have evaluated the inhibition of these compounds *in vitro* against *Ng*CysE. Experimental screening led to the identification of first inhibitors of *Ng*CysE, with the most potent compound displaying inhibition in the low micromolar range.

The work presented in this chapter is shown in a publication format as a manuscript prepared for submission to the journal; *Antimicrobial Agents and Chemotherapy* (American Society for Microbiology). Included in this chapter is a Future Directions section (4.7), which itself is not part of the manuscript, but has been included to discuss ongoing and future work directly related to the project; these results will be included in the future publications. Supplementary information is available in Appendix B. Citation for this work is as follows:

Oldham, K. E. A., Jiao, W., & Hicks, J. L. (2024). *Identification of novel inhibitors targeting serine acetyltransferase from Neisseria gonorrhoeae.*

(Prepared for submission)

4.1.1 Author contributions

I led the experimental work for this research, as first author, which involved enzyme purification, optimisation of enzyme kinetic assays, experimental compound

screening, analysis of compound 2 interactions with *Ng*CysE. Furthermore, I prepared figures as well as manuscript drafting and editing. Wanting Jiao conducted the virtual inhibitor screening, including preparation and screening of the ligand libraries against the *Ng*CysE model, as well as figure preparation, manuscript drafting and editing. Joanna Hicks conceptualised research, provided funding for project, figure preparation as well as manuscript drafting and editing. The authorship contribution form can be found in Appendix D.

Identification of novel inhibitors targeting serine acetyltransferase from *Neisseria gonorrhoeae*

Keely E.A. Oldham[1,2], Wanting Jiao[3], Joanna L. Hicks[1#]

[1]*Te Huataki Waiora School of Health, University of Waikato, Hamilton 3216, New Zealand*

[2]*Te Aka Mātuatua School of Science, University of Waikato, Hamilton 3216, New Zealand*

[3]*The Ferrier Research Institute, Victoria University of Wellington, Wellington 6140, New Zealand*

[#]Address correspondence to Joanna Hicks, joanna.hicks@waikato.ac.nz

KEYWORDS

Serine acetyltransferase, CysE, antibiotic resistance, cysteine biosynthesis, inhibitor screening.

4.2 Abstract

Neisseria gonorrhoeae is an obligate human pathogen and the etiological agent of the sexually transmitted infection, gonorrhoea. The rapid emergence of extensively antimicrobial-resistant strains, including those resistant to all frontline antibiotics, has led to *N. gonorrhoeae* being labelled a priority pathogen by the World Health Organisation highlighting the need for new antimicrobial treatments. Targeting the *de novo* cysteine biosynthesis pathway has been identified as a promising avenue for developing new antimicrobials against *N. gonorrhoeae*. Here, we have conducted virtual inhibitor screening of commercially available compound libraries against the key cysteine biosynthetic enzyme, serine acetyltransferase from *N. gonorrhoeae* (*Ng*CysE, EC 2.3.1.30). We have identified a hit compound with an IC_{50} of 13 µM. Future research will involve inhibitor optimisation and development as a promising antimicrobial target for *N. gonorrhoeae.*

4.3 Introduction

Antimicrobial resistance is an ever-growing global problem and has been declared a public health crisis. *Neisseria gonorrhoeae* (gonococcus) is a Gram-negative bacterium and is an obligate human pathogen, causing the sexually transmitted infection (STI), gonorrhoea. Globally, there are approximately 87 million cases annually, making it the second most common bacterial STI, after chlamydia (Wi et

al., 2017). Concerningly, *N. gonorrhoeae* can readily acquire antimicrobial resistance, with documented resistance emerging for every class of antibiotic that has been used for its treatment since the introduction of sulphonamides in the 1930s (Unemo, Golparian, & Hellmark, 2014). In 2017, *N. gonorrhoeae* was labelled as a priority pathogen by the World Health Organisation, (WHO, 2017) and in the absence of an effective vaccine (Murray et al., 2022), coupled with the increasing prevalence of antibiotic-resistant strains, there is an urgent need for new antimicrobial treatments.

An emerging theme in the development of new antimicrobials is the targeting of key metabolic pathways that are important for infection and virulence. As such, there is expanding interest in targeting amino acid biosynthetic pathways in bacteria. One such pathway is the *de novo* biosynthesis of the amino acid L-cysteine. This pathway is well-conserved across bacteria and plants but is notably absent in mammals (Kredich, 2008). Alongside being an important amino acid in protein molecules, cysteine is a versatile thiol donor and used for the synthesis of thiol-containing metabolites, such as glutathione and thioredoxin, both of which are important in regulating the cell redox state and for protection from oxidative stress encountered during infection (Kredich, 2008). Cysteine is synthesised via a twostep reaction by the cysteine biosynthetic enzymes, serine *O*-acetyltransferase (SAT/CysE, EC 2.3.1.30) and *O*-acetylserine sulphydrylase (OASS-A/OASS-B; CysK/CysM, EC 4.2.99.8) (Benoni et al., 2017; Kredich & Tomkins, 1966). CysE catalyses the acetylation of serine using acetyl-CoA to produce *O*-acetylserine which is subsequently condensed with sulphide by CysK or thiosulphate in the case of CysM, to form L-cysteine (Figure 4.1).

The *de novo* synthesis of cysteine is tightly regulated at the protein level which is primarily through feedback inhibition of CysE by cysteine, as high concentrations can lead to damaging Fenton chemistry (Park & Imlay, 2003). Additionally, cysteine synthesis is regulated with regard to sulphur availability through the formation of the cysteine synthase complex (CSC). The CSC is formed the reversible association of one CysE hexamer and two CysK dimers formed through the CysE C-terminal tail binding into and occluding the CysK active site, enhancing CysE and abolishing CysK activity. CSC formation is favoured under low sulphide

concentrations (Kredich, Becker, & Tomkins, 1969) which allows *O*-acetylserine to accumulate and dissociate the CSC (Benoni et al., 2017).

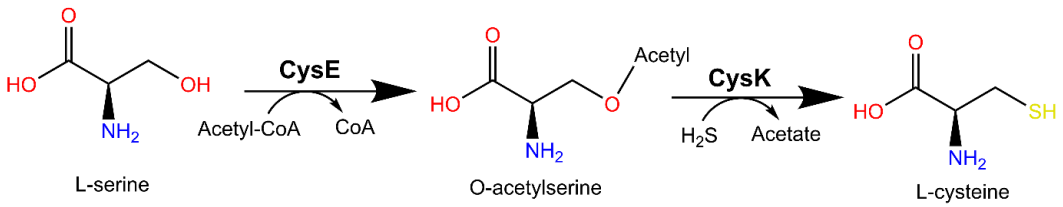


Figure 4.1. *De novo* cysteine biosynthesis reaction in *Neisseria gonorrhoeae.* CysE, serine *O*-acetyltransferase; CysK, *O*-acetylserine sulphydrylase.

The absence of the cysteine biosynthesis pathway in humans makes it ideal for antimicrobial targeting in reducing the chance of off-target effects (Griffith, 1987). Additionally, targeting of CysE has been reviewed (Hicks, Oldham, McGarvie, & Walker, 2022) and reported for notable bacterial pathogens including *Escherichia coli* (Toyomoto et al., 2023), *Salmonella typhimurium* (Magalhães et al., 2020) *Staphylococcus aureus* (Chen et al., 2019) and *Mycobacterium tuberculosis* (Gupta & Gupta, 2021). The CysE from *N. gonorrhoeae* (*Ng*CysE) has been shown to be differentially expressed during the stringent response (Talà et al., 2023) and has been identified as an essential gene in *N. gonorrhoeae* (Remmele et al., 2014), thus we propose *Ng*CysE is a promising new target for antimicrobial development. Previously, we have kinetically and structurally characterised *Ng*CysE (PDB 6WYE and 7RA4) demonstrating that it has comparable activity to other bacterial CysE homologues and possesses conserved structural fold and active site residues (Oldham et al., 2022). Here, we have used structure-based virtual inhibitor screening of commercial compound libraries against *Ng*CysE and tested *in vitro* 28 hit compounds with one compound (compound 2) identified as a potent inhibitor of *Ng*CysE.

4.4 Results

4.4.1 CysE is highly conserved across *Neisseria gonorrhoeae* strains

Conservation of active site residues is critical for ensuring broad spectrum activity against different *N. gonorrhoeae* isolates. Analysis of the presence and conservation of CysE (*cysE*) amongst *N. gonorrhoeae* strains was conducted via BLAST search using *cysE* from *N. gonorrhoeae* MS11 as a query against all available *N. gonorrhoeae* genomes in the NCBI database. *cysE* was shown to be

highly conserved for *N. gonorrhoeae* (99.88-100% sequence identity) over the length of the sequence (819 nucleotides, 272 residues) with the majority of strains having an identical isoform (156/157 genomes) (Figure B.1). To explore the distribution of CysE amongst *N. gonorrhoeae* isolates, we expanded our search to include all available *N. gonorrhoeae* genomes in the PubMLST database. A *cysE* gene (NEIS0501) was annotated in all isolate genomes (19,729 genomes) and is well-conserved, with 97.7-100% similarity reported. Furthermore, comparison of these isolates show that the majority genomes had an annotated *cysE* identical to the MS11 *cysE* query (18,978/19,729, 96% of genomes had an identical copy). Overall *cysE* is strictly conserved across *N. gonorrhoeae* reference strains and circulating isolates, thus making it a suitable target for antimicrobial applications.

4.4.2 *Ng*CysE active site model with both substrates bound

The crystal structure of *Ng*CysE (PDB 6WYE) contains L-malate in the serine binding site and lacks acetyl-CoA. To examine substrate-enzyme interactions, we generated an active site model with both substrates bound. The crystal structures of *Yersinia pestis* CysE (PDB 3GVD) and *Haemophilus influenzae* CysE (PDB 1SST) contain the inhibitor cysteine and the product co-enzyme A in the active site, respectively. With the structural similarity between serine and cysteine these two structures were used as the starting point for positioning serine and acetyl-CoA into the active site of *Ng*CysE (more detail in the Methods section). The initial active site model of *Ng*CysE with both serine and acetyl-CoA placed was then further optimised through a series of enzyme minimisation, refinement, and quantum mechanics/molecular mechanics (QM/MM) optimisation.

The active site model (Figure 4.2A) shows that both substrates can bind simultaneously. The nucleophilic oxygen atom of serine is 3.9 Å away from the acetyl carbon of acetyl-CoA, positioning it for the reaction. Acetyl-CoA primarily interacts with the active site via polar interactions from its phosphate groups (Figure 4.2). The phosphate forms salt bridge interactions with two Lys246 residues from the two chains forming the active site. The diphosphate group establishes a salt bridge interaction with Lys223. Additionally, acetyl-CoA forms hydrogen bond (H-bond) interactions with the backbones of residues Ala226, Ala208, and Ala249. The binding of serine is stabilised by interactions between its backbone and charged

residues Arg196, Asp96, and Asp161. The side chain of serine interacts with two histidine residues, His197 and His162. The current active site model suggests that His162 will act to deprotonate serine for the reaction.

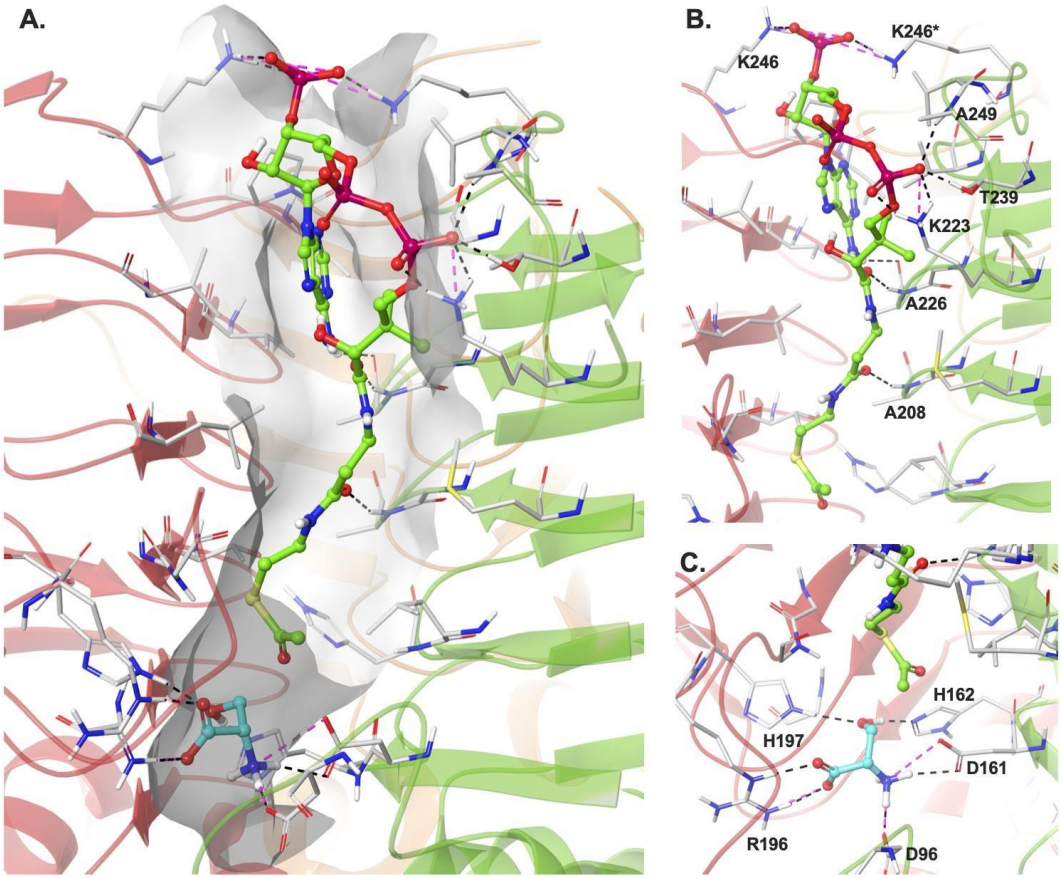


Figure 4.2 Active site model of *Ng*CysE with both serine and acetyl-CoA bound. (A) Binding sites and relative positions of serine (cyan carbon) and acetyl-CoA (green/pink carbon). (B) and (C) binding sites for acetyl-CoA and serine, respectively. Residues involved in polar interactions are labelled. H-bond interactions are displayed as black dashed lines and salt bridges are displayed as magenta dashed lines.

4.4.3 Hit inhibitor screening and characterisation

We employed a structure-based virtual screening approach to search for potential *Ng*CysE inhibitors. The optimised active site model with both substrates bound was used as the receptor for virtual screening. Compounds with drug-like properties were obtained from the ZINC database (Sterling & Irwin, 2015) which, after preparation, contained a total of ~10.5 million compounds. The drug-like compound library was screened against the substrate-bound crystal structure model of *Ng*CysE, with the acetyl-CoA binding site being the primary target binding site for screening.

The top 1000 poses from the screening calculation were examined for the enzyme-ligand interactions. Compounds showing interactions to the backbones of Ala208 and Ala226, and additional interactions to residues in the acetyl-CoA binding site were selected for further molecular mechanics with generalised Born and surface area calculations (MM-GBSA) (54 compounds). During the MM-GBSA calculations, active site residues within 5 Å of the docked compound were refined to optimise the enzyme-ligand interactions. A binding free energy was also computed for each compound. Following the MM-GBSA calculation, compounds that had a predicted binding free energy below -50 kcal/mol and retained the interactions with the backbones of Ala208 and Ala226, were selected for validation using short (2 ns) molecular dynamics (MD) simulations (23 compounds). The MD trajectories of the ligand-enzyme complex were examined and compounds that maintained their interactions during the MD simulations were selected for testing (nine compounds).

In addition to the drug-like compound library, a second library containing large compounds with a molecular weight greater than 500 Daltons was screened. The compounds were retrieved from the ZINC database (Sterling & Irwin, 2015) and prepared for docking. The prepared library contained ~1.3 million compounds. The target receptor conformation used for screening the big-compound library was a representative active site conformation obtained from a 5 μs MD simulation of the *Ng*CysE enzyme. The acetyl-CoA site was the target screening site. The top 1000 compounds were examined using criteria similar to those mentioned above; 152 compounds underwent MM-GBSA calculation, of which 82 maintained backbone interactions with the alanine residues and were predicted to have binding free energies below -60 kcal/mol. The 82 compounds were clustered into 36 groups based on their structures, and one representative compound from each group was recommended for further testing. Of the 45 recommended compounds, 28 could be purchased at the time of the study. The details of hit compounds experimentally tested can be found in (Table B.1).

4.4.4 *In vitro* screening and characterisation

Hit compounds identified from virtual inhibitor screening were experimentally evaluated by measuring the inhibition of recombinant *Ng*CysE activity.

Recombinant *Ng*CysE was expressed and purified following published methods (Oldham et al., 2022). *Ng*CysE was expressed with a N-terminal HexaHistag and purified using both immobilised affinity chromatography and size exclusion chromatography. Previously, we had characterised *Ng*CysE using a direct 232 nm assay to determine the Michaelis-Menten kinetic parameters for substrates serine and acetyl-CoA (Oldham et al., 2022). Given the high absorbance of aromatic compounds at this wavelength, *Ng*CysE inhibitor screening was performed using a plate-based coupled assay detecting production of CoA using 5,5'-Dithio-bis-(2-nitrobenzoic acid) (DTNB) through measuring production of 2-nitro-5-thiobenzoate (TNB) adapted from (Magalhães et al., 2020). *Ng*CysE activity was determined from a co-enzyme A DTNB standard curve (Figure B.2).

For *in vitro* screening, 28 compounds were ordered from commercial vendors and 1 mM stocks were prepared for screening in 100% dimethyl sulfoxide (DMSO). Initial screening at 100 µM in the presence of 1 mM serine and 0.15 mM acetyl-CoA was conducted to identify inhibition and check solubility (Figure 4.3, Table C.1). Compounds that showed inhibition below 50% and did not precipitate in the assay reaction were subjected to IC_{50} screening (compounds 1, 2, 3, 4 and 5). Upon further testing, compounds 1, 3, 4 and 5 did not display dose-dependent inhibition and therefore were classified as false positives from the original screen (data not shown). Compound 2, however, demonstrated inhibition (20.6% remaining activity at 100 µM) and collection of a dose response curve produced an IC_{50} of 13.9 ± 4.3 µM (R^2 = 0.9924) (Figure 4.4) making it the most potent hit compound.

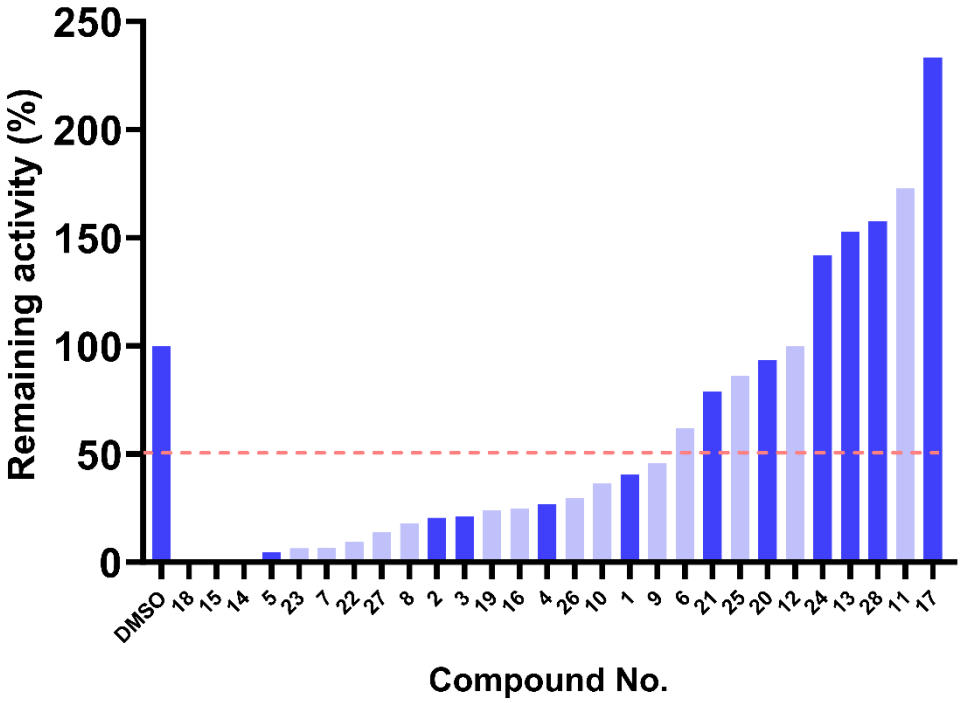


Figure 4.3 100 µM activity screen of library compounds against *Ng*CysE. Remaining activity (%) is calculated as rate with inhibitor present divided by rate in presence of DMSO (V_i/V_o) multiplied by 100. Soluble and insoluble compounds are highlighted in, dark and light blue, respectively. Dashed pink line corresponds with 50% activity cut-off for downstream testing. Compounds 14, 15 and 18 have negative rates and for display purposes are represented as zero. Well containing DMSO only treated as positive control (100% activity).

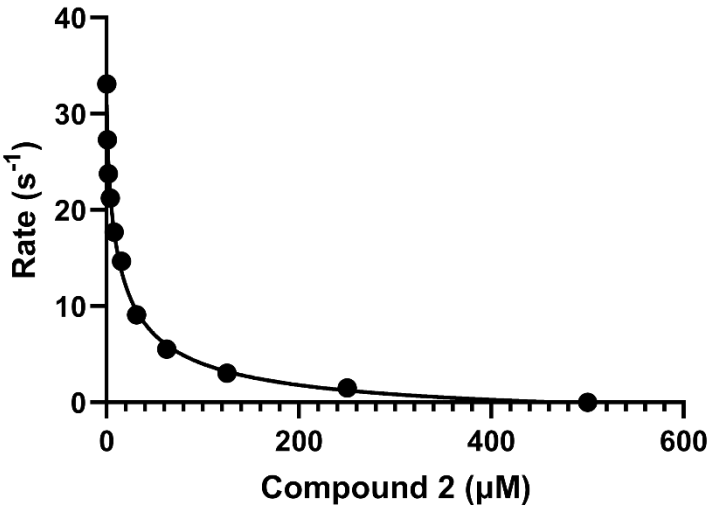


Figure 4.4 IC_{50} dose response curve for Compound 2. *Ng*CysE activity measured in presence of 0.1 mM acetyl-CoA and 0.5 mM serine. Rate (s^{-1}) was calculated as per hexamer. Datapoints are from a single reaction.

4.4.5 Docking Interactions for Compound 2

Compound 2 (1-(6-ethoxy-4-methyl-2-quinazolinyl)-N-[2-(1H-imidazol-1yl)ethyl]-3-piperidinecarboxamide), consists of; an imidazole ring, with an amide linker connected to a piperidine ring, which itself is connected to a quinazoline ring with an ether attached (Figure 4.5). The compound does not violate any Lipinski's rules (Table 4.1) and has qualities consistent with a drug-like compound (Lipinski, 2000).

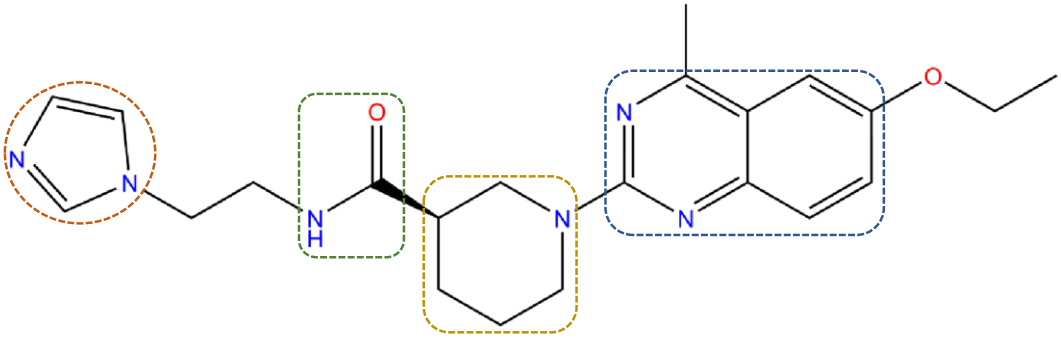


Figure 4.5 Structure of Compound 2(1-(6-ethoxy-4-methyl-2-quinazolinyl)-N-[2(1H-imidazol-1-yl)ethyl]-3-piperidinecarboxamide). Functional groups imidazole, piperidine and quinazoline are highlighted in orange, yellow, and blue, respectively. Figure created using ChemDraw Prime.

Table 4.1 Compound 2 chemical properties. tPSA, polar surface area (Å^2); logP, partition coefficient.

Chemical Formula	Net charge	H-bond donors	H-bond acceptors	Molecular Weight (g/mol)	logP	tPSA	Rotatable bonds	Apolar desolvation	Polar desolvation
$C_{22}H_{28}N_6O_2$	0	1	7	408.5	2.566	85	7	9.87	-17.44

Examination of *Ng*CysE poses with compound 2 docked from virtual inhibitor screening, shows compound 2 forms notable active site interactions with the serine and acetyl-CoA binding sites. There are four H-bond interactions formed between compound 2 and serine binding site residue Asp161 (Chain C) and with backbone interactions with acetyl-CoA binding residues Ala208 (Chain C), Ala226 (Chain C), and Gly188 (Chain A), and is complimented with a number of hydrophobic interactions from the acetyl-CoA site with the piperidine and quinazoline ring (Figure 4.6). Additionally, these interacting serine and acetyl-CoA site residues are strictly conserved not only for *Neisseria* but also Proteobacteria (Figure 4.7).

Furthermore, compound 2 does not share a high degree of similarity with published CysE inhibitors based on Tanimoto similarity scores (Figure 4.8) (scores lower than 0.7 (Bajusz, Rácz, & Héberger, 2015)), making compound 2 a novel CysE inhibitor.

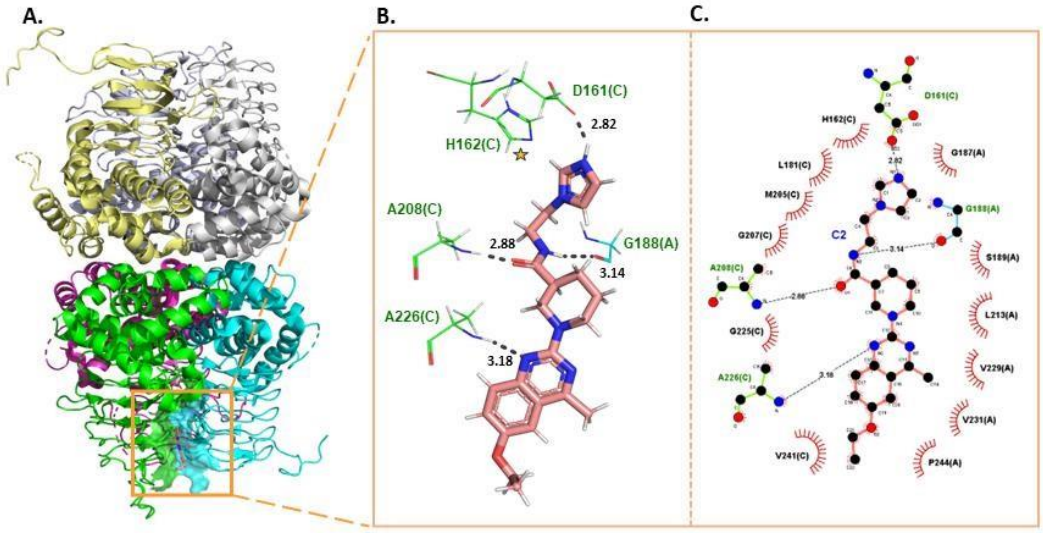


Figure 4.6 Docked interactions of compound 2 binding to *Ng*CysE. (A) *Ng*CysE hexamer (6WYE) with compound 2 forms hydrogen bond interactions with *Ng*CysE serine and acetyl-CoA binding pockets (Chain A and Chain C, coloured cyan and green, respectively). (B) Hydrogen bond interactions (grey) between compound 2 (pink sticks) and *Ng*CysE residues (green and cyan wires). (C) Two-dimensional LigPlot schematic of H-bonds and hydrophobic interactions between compound 2 and *Ng*CysE residues (bond lengths not to scale). H-bonds are shown as black dashes and hydrophobic interactions are represented as red radiating lines. Catalytic His162 (H162) is indicated with yellow star. All bond lengths are reported in Angstroms (Å). Figures created using PyMOL and LigPlot^{+} (Laskowski & Swindells, 2011).

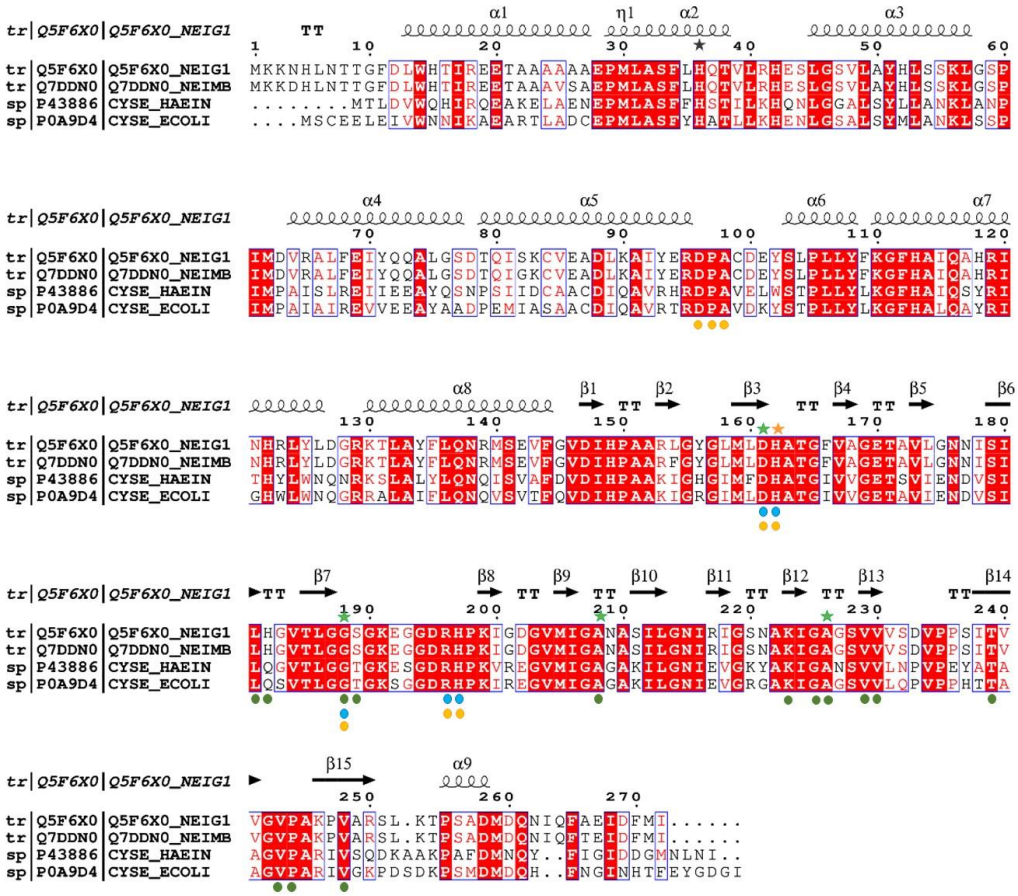


Figure 4.7 Multiple sequence alignment of bacterial CysE isoforms. Serine, acetyl-CoA and cysteine binding residues are annotated as green, blue and orange dots, respectively. Compound 2 hydrogen bond interactions are annotated by green stars. Active site catalytic His162 is highlighted as orange star. Sequences: NEIG1, *Neisseria gonorrhoeae* (PDB 6WYE); NEIMB, *Neisseria meningitidis* (no crystal structure); HAEIN, *Haemophilus influenzae* (PDB 1SST); ECOLI, *Escherichia coli* (PDB 1T3D).

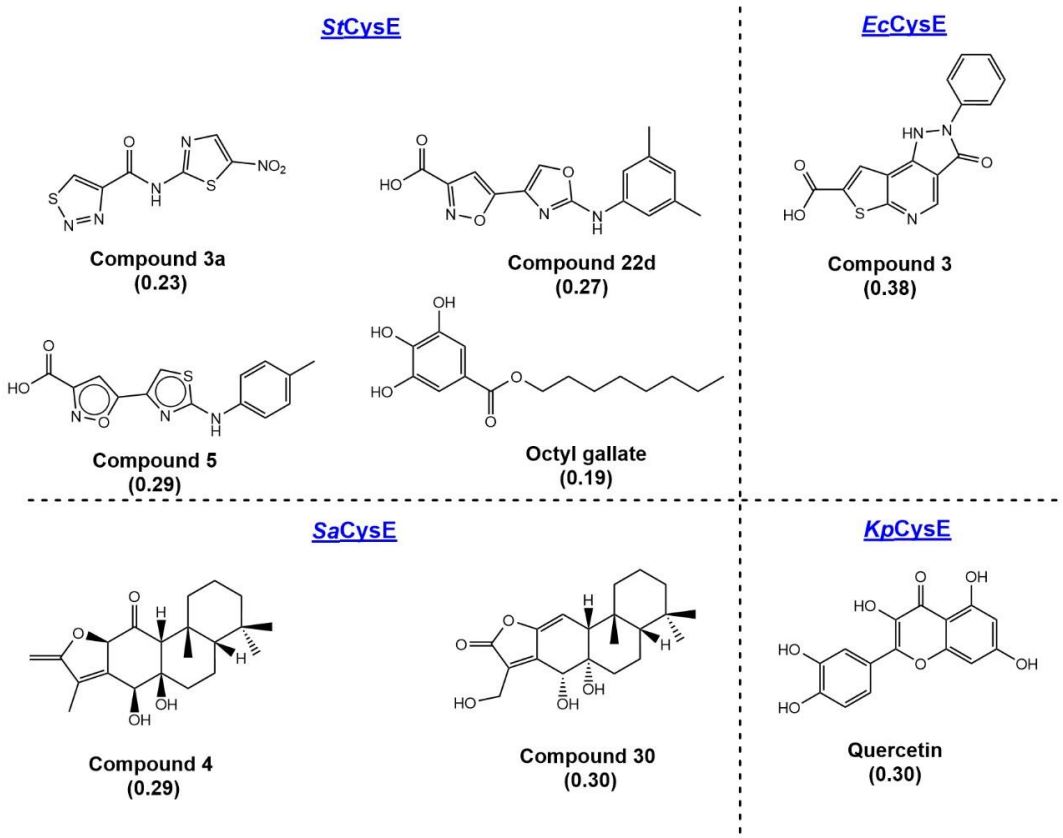


Figure 4.8 Characterised inhibitors against bacterial CysE homologues. Tanimoto fingerprints co-efficient similarity scores of reported CysE inhibitors compared to compound 2 are shown in brackets. Tanimoto scores were calculated using RDKit (2023).

4.5 Discussion

Antimicrobial-resistant *N. gonorrhoeae* is a global burden on human health, and in the absence of an effective vaccine, antimicrobials remain the front-line tool for managing gonococcal disease. The targeting of the *de novo* cysteine biosynthesis pathway has been highlighted as an avenue for development of antimicrobial compounds and adjuvants given its role in mitigating oxidative stress and the absence of the pathway in humans. Here we report the first inhibitor of *Ng*CysE and demonstrate that structure-based virtual inhibitor screening is a successful strategy for identifying inhibitors. After computationally screening two libraries of 1.3 and 10.5 million compounds from the ZINC15 database, 45 compounds were identified that met our criteria *in silico*, and of the 28 compounds tested experimentally, the best compound to inhibit *Ng*CysE activity *in vitro* was compound 2. Docking analysis revealed notable H-bond and hydrophobic interactions between

compound 2 and both *Ng*CysE serine and acetyl-CoA binding pockets along the length of the compound. Dose-response analysis of compound 2 showed it to be a promising inhibitor with an IC_{50} of 13.9 µM.

The use of both virtual and experimental testing is needed for the rapid identification of novel inhibitors and has proved to be a successful strategy for identifying bacterial CysE inhibitors. The inhibition observed for compound 2 is in the low micromolar range, making it one of the better CysE inhibitors reported to date (Hicks et al., 2022). Although inhibitors that exhibit sub-micromolar inhibition have been reported, ensuring that the compound is membrane permeable remains a challenge, particularly given the cytoplasmic nature of CysE isoforms, including *Ng*CysE, (Magalhães et al., 2021) and is a notorious problem for Gram-negative bacteria (Maher & Hassan, 2023). Therefore, determining the MIC of compound 2 against *N. gonorrhoeae* is needed to ensure the compound is bioactive and thus a viable antimicrobial.

As highlighted previously (Hicks et al., 2022), *in vivo* validation of CysE inhibitors has often been limited in their ability to disrupt bacterial growth. Only recently, Toyomoto and colleagues showed CysE inhibition in *E. coli* could not only disrupt bacterial growth, but also deplete intracellular concentrations of thiol-metabolites, as well as increase sensitivity to antibiotics and oxidative stress (Toyomoto et al., 2023). This work highlights the multi-faceted role CysE serves within bacteria, indicating that not only will *de novo* cysteine biosynthesis be disrupted by inhibition, but also resistance to oxidative stress, and existing resistance to known antibiotics.

A caveat of antimicrobial targeting of *de novo* cysteine biosynthesis is the possibility of resistance via up-regulation of alternative sulphur acquisition and assimilation pathways to obtain cysteine (Toyomoto et al., 2023). Sulphur acquisition in *N. gonorrhoeae* is unique in that it possesses a sulphate/thiosulphate ABC transporter (Hicks & Mullholland, 2018), thiol amino acid transporters (methionine (Semchenko, Day, & Seib, 2017), cysteine, and cystine (Bulut et al., 2012; Semchenko et al., 2017)), but lacks importers for thiol compounds taurine or glutathione (Seib et al., 2006). Additionally, glutathione is maintained at comparatively high levels in *N. gonorrhoeae* (15 mM vs. 5 mM *E. coli*) (Archibald & Duong, 1986), and given the absence of glutathione importers or the glutathione

recycler γ-glutamyl transpeptidase (Takahashi & Watanabe, 2005), *N. gonorrhoeae* is reliant on *de novo* biosynthesis of glutathione and would rely on cysteine to maintain these high glutathione concentrations. Given this lack of pathway redundancy and high need for antioxidant defence during infection, and the essential nature of *cysE*, *N. gonorrhoeae* is unlikely to develop resistance to CysE inhibitor making it an ideal organism for *de novo* cysteine inhibition.

Cysteine has also been shown to contribute to virulence for *N. meningitidis*. The sulphate reduction arm of the *de novo* cysteine biosynthesis pathway is differentially expressed upon exposure to epithelial cells for *N. meningitidis* (Dietrich et al., 2003; Grifantini et al., 2002; Joseph et al., 2010). Given the large degree of sequence conservation between CysE homologues from *N. gonorrhoeae* and *N. meningitidis* (97.8 % sequence identity, six residue differences, Figure 4.7) and previous research demonstrating cross-species inhibition for CysE inhibitors (Magalhães et al., 2020; Toyomoto et al., 2023) hit compounds reported here could be screened against *Nm*CysE. Considering the essential nature of *cysE* in *N. meningitidis* (Capel et al., 2016; Muir et al., 2020), *Nm*CysE is a potential target for antimicrobial discovery.

In summary, we have demonstrated that structure-based virtual inhibitor screening is a promising avenue for identifying inhibitors of *Ng*CysE. We report compound 2 to show inhibition of *Ng*CysE in the low micromolar range making it one of the most potent CysE inhibitors to date. Furthermore, given the low similarity to published bacterial CysE inhibitors compound 2 is a novel CysE inhibitor. Given the essential nature of *cysE* in *N. gonorrhoeae*, compound 2 is a promising candidate for antimicrobial development against this pathogen.

4.6 Materials and Methods

4.6.1 Bioinformatic analysis of *Ng*CysE

Investigation of the conservation and distribution of CysE sequences in *N. gonorrhoeae* was conducted through a BLASTN search of the complete *N. gonorrhoeae* genomes (taxid: 485) held by NCBI (“Complete prokaryote Genome Database”) using *Ng*CysE nucleotide sequence from *N. gonorrhoeae* strain MS11 as a query (RefSeq Accession; WP_025457874.1, locus_tag

NGFG_RS07905, NC_022240.1:1458601-1459419). Determining CysE presence in *N. gonorrhoeae* isolates (24,173 isolates) held in the PubMLST (accessed Jan 2024) (Jolley, Bray, & Maiden, 2018) was conducted via searching the CysE sequence from *N. gonorrhoeae* strain MS11 as a query (RefSeq Accession; WP_003989278.1, (NC_022240.1:1458601-1459419), NEIS0501) using the BLASTN search plug-in.

4.6.2 Substrate-bound structural model of *Ng*CysE

A hexamer of *Ng*CysE was generated from symmetry operations on the crystal structure of *Ng*CysE (PDB 6WYE). The malate molecules were deleted. Crystal structures of *Yersinia pestis* CysE (PDB 3GVD), which contains cysteine in the serine site, and *Haemophilus influenzae* CysE (PDB 1SST), which contains co-enzyme A, were aligned with the *Ng*CysE crystal structure at chains A and C (*Ng*CysE chain names). The cysteine and co-enzyme A molecules binding in the A/C interface were extracted from 3GVD and 1SST and merged with the *Ng*CysE crystal structure. The cysteine ligand in the active site was then mutated to serine *in silico* in Maestro from Schrödinger Suite (Schrödinger, 2020e). Thus, a structure of *Ng*CysE with CoA and serine bound in the active site was generated. This merged structure was then prepared for modelling using the protein preparation wizard in the Schrödinger Suite (Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2013; Schrödinger, 2020g). During protein preparation, hydrogen atoms were added, ionisation states for the ligands were generated using Epik (Greenwood, Calkins, Sullivan, & Shelley, 2010; Schrödinger, 2020b; Shelley et al., 2007) for pH 7±2, and all water molecules were deleted. Hydrogen bond pairs were optimised, and protonation states of ionisable residues were assigned using PROPKA (Li, Robertson, & Jensen, 2005; Sastry et al., 2013) for pH 8.0. The structure was then subjected to energy minimisation, converging heavy atoms to an RMSD of 0.3 Å. An acetyl group was then built onto the CoA molecule, followed by another round of restrained minimisation, converging heavy atoms to an RMSD of 0.3 Å.

The active site model of *Ng*CysE containing serine and acetyl-CoA in the active site in the A/C interface was further optimised using Prime refinement (Jacobson, Friesner, Xiang, & Honig, 2002; Jacobson et al., 2004; Schrödinger, 2020f). Atoms within 8 Å of the serine and acetyl-CoA substrates were minimised to sample

conformational changes around the serine and acetyl-CoA binding site and to better accommodate the ligand molecules that were absent in the initial crystal structure. Finally, QM/MM optimisation calculation was set up using QSite (Murphy, Philipp, & Friesner, 2000) from the Schrödinger Suite (Schrödinger, 2020h), to search for catalytic-relevant conformations (positionings of the serine and acetyl group in relation to the catalytic dyad). The QM region was defined to include side chains of His162, Asp161, serine substrate, and part of the acetyl-CoA molecule defined by inserting a hydrogen cap between C6P and C7P on acetyl-CoA. The resultant QM region contained 51 atoms. The QM optimisation was set up using the DFT-B3LYP method and lacvp* basis set. The MM region was minimised using the Truncated Newton algorithm with a maximum of 1000 cycles, and the QM region was minimised with a maximum of 5000 iterations.

4.6.3 Molecular dynamics simulation of *Ng*CysE

The crystal structure of *Ng*CysE contains missing atoms in disordered regions. Before MD simulations could be set up, a complete structure containing residues 1 to 264 was generated by building in the missing residues and side chains in COOT (Emsley, Lohkamp, Scott, & Cowtan, 2010), fitting to the electron density map. The complete hexamer was then generated based on crystal symmetry. The malate molecules were deleted. Crystal structures of *Yersinia pestis* CysE (PDB 3GVD) and *Haemophilus influenzae* CysE (PDB 1SST) were aligned to the top and bottom trimers of *Ng*CysE (as there appears to be a degree of rotation relative to the top and bottom trimers in *Ng*CysE compared to the other two crystal structures). Then, all six cysteine and the four CoA molecules in the "open" active sites of *Ng*CysE were extracted from 3GVD and 1SST and merged with the *Ng*CysE structure. The cysteine ligands in the active sites were then mutated to serine *in silico*. Thus, a structure of the complete hexamer of *Ng*CysE with four CoA and six serine bound in the active site was generated.

This merged structure was prepared for modelling using the protein preparation wizard (Sastry et al., 2013; Schrödinger, 2020g). Hydrogen atoms were added, and protonation states of ligands were assigned using Epik for a pH range of 7±2. The amino group of serine ligands were fixed to have a +1 charge, and the carboxylate terminal group of Gln264 residues was added. To remove steric clashes between

merged substrates and the binding site residues, atoms on CoA and residues within 5 Å of the substrates were minimised using Prime energy minimisation calculation. Then, hydrogen bonds were optimised, and protonation states of ionisable residues were assigned using PROPKA for pH 8.0. It should be noted that His162 was assigned a protonation state of HIS. An acetyl group was built onto each co-enzyme A molecule, followed by another round of Prime minimisation for acetyl-CoA and residues within 5 Å of acetyl-CoA and serine.

In this new structure, there were two types of active site conformations: fully occupied active site (with both serine and acetyl-CoA) and partially occupied closed active site (with serine present but the acetyl-CoA site occluded by the closing C-terminal tail). The new complete structure with substrates bound was used to set up MD simulations. The MD system was built in Desmond (Bowers et al., 2006; Schrödinger, 2020a) by adding TIP3P explicit water molecules to an orthorhombic box. The box size was set by having a buffer region of 10 Å in each of the x, y, z directions. The box volume was minimised by rotating the enzyme molecule. Ions were added to neutralise the system, and NaCl salt was added at a concentration of 0.15 M. The force field used was OPLS_2005. MD simulation was conducted using Desmond (Bowers et al., 2006; Schrödinger, 2020a). The system was first minimised for 100 ps, followed by NPT ensemble MD calculations at a temperature of 310 K and a pressure of 1.01325 bar. Langevin thermostat and barostat were used. The MD simulation was conducted for 5 μs.

4.6.4 Ligand Library Preparation

Drug-like compounds with neutral charges were retrieved from the ZINC15 database (Sterling & Irwin, 2015). The 3D structures of these compounds, possessing drug-like properties, devoid of reactive groups, charge-neutral at pH 7, and available for purchase, were downloaded from the ZINC15 database (accessed on 28-08-2019). The initially downloaded library comprised approximately 7 million compounds. This compound set was subsequently filtered using Pan-assay interference compounds (PAINS) filters (Baell & Holloway, 2010) and prepared for modelling in LigPrep (Schrödinger, 2020d), using the OPLS3 force field. Possible ionisation states were generated for pH 7±1 using Epik, and potential tautomers were generated while retaining the specified chiralities in the compounds.

The prepared library now consists of 10,479,301 compounds. For the large compound library, 3D structures of compounds with a molecular weight greater than or equal to 500 Daltons, a logP value above 4, charge-neutral at pH 7, no reactive groups, and available for purchase, were downloaded from the ZINC15 database (accessed on 2021-06-21) (Sterling & Irwin, 2015). The initial library includes approximately 1 million compounds. The set of compounds was already filtered by PAINS filters upon download (Baell & Holloway, 2010). The library was prepared using LigPrep with the OPLS4 force field. Possible states at pH 7±1 were generated using Epik; tautomers were generated while retaining chiralities within the 3D structures. The prepared library contains 1,294,740 compounds.

4.6.5 Receptor grid generation

The substrate-bound active site model was used to generate the receptor grid for virtual screening. Firstly, the acetyl-CoA and serine substrates were removed. Two receptor grids were generated with slightly different positions for the center of the grid. For the first receptor grid, the center of the grid was defined as the centroid of residues 246, 231, 229, and 244 from chain A, and residues 246, 223, 205, 226, and 208 from chain C. For the second receptor grid, the center of the grid was defined as the centroid of residues 246, 231, and 244 from chain A, and residues 246 and 248 from chain C. The size of both grids was defined by allowing docking of ligands with a length <= 20 Å. For the first grid, rotation of the hydroxyl groups on Thr239 and Ser228 from chain C was allowed. For the second grid, rotation of the hydroxyl groups on Ser189 from chain A, and Ser228 and Thr239 from chain C was allowed. The receptor grids were generated using Glide (Friesner et al., 2004; Friesner et al., 2006; Halgren et al., 2004) from the Schrödinger Suite (Schrödinger, 2020c).

During the 5 μs MD simulation for the full-length *Ng*CysE structure, the acetyl-CoA molecule bound in the interface between chains E and F remained stable; thus, this binding site was used for virtual screening of the large-compound library. A representative conformation (frame at 3.64 μs from the full trajectory) was extracted from the equilibrated time period of the MD simulation (3-5 μs) to be used as the receptor conformation for virtual screening. This conformation was prepared for receptor generation using the protein preparation wizard (Sastry et al., 2013; Schrödinger, 2020g). All water molecules were deleted, and hydrogen bond

interactions as well as protonation states of ionisable residues were assigned using PROPKA for pH 8.0. Restrained minimisation of the structure was then conducted to remove any steric clashes. A receptor grid was generated centering at the acetyl-CoA binding site between chains E and F. All other acetyl-CoA and serine ligands were deleted. The centre of the grid was defined by the position of acetyl-CoA in the E/F interface. The size of the grid was determined by allowing docking of ligands with sizes similar to acetyl-CoA. Rotatable groups were allowed for side chains of Ser189 from chain F and Thr254 from chain E.

4.6.6 Virtual screening

The prepared compound library was screened against the receptor grids generated above using the virtual screening workflow in Schrödinger Suite. Virtual screening (VS) was conducted in three docking stages with increasing precision, HTVS (high throughput VS), SP (standard precision), and XP (extra precision). Top 10% ranked compounds from the HTVS stage were subject to SP docking, and the best 10% compounds from the SP stage were then subject to the final XP docking state. One pose was written out for each ligand, and the top ranked 1000 compounds were written out for evaluation.

4.6.7 MM-GBSA calculation and MD validation

The docked poses of selected compounds, after examination of the virtual screening results, were further evaluated using MM-GBSA (molecular mechanics with generalised Born and surface area) calculations in Prime (Jacobson et al., 2002; Jacobson et al., 2004; Schrödinger, 2020f). MM-GBSA binding energies, representing approximate free energies of binding, were calculated as the difference between the energy of the ligand-enzyme complex and the sum of energies of the free ligand and free enzyme. A more negative value indicates a predicted stronger binding. MM-GBSA calculations were conducted using the OPLS4 force field and the VSGB solvation model. The binding poses were refined, and active site residues within 5 Å of the docked compound were minimised to optimise interactions.

After further examination of the MM-GBSA results, compounds that retained interactions with the enzyme and showed favourable predicted binding energies were selected for additional validation through short MD simulations (2 ns). The

MD simulation system for each compound was constructed based on the MMGBSA optimised binding pose, and explicit SPC water molecules were added to orthorhombic boxes. The box size was determined by establishing a buffer region of 10 Å in each of the x, y, and z directions. The box volume was minimized by rotating the enzyme-ligand complex. Na^+ and Cl^- ions were incorporated to neutralise the system, and NaCl salt was added at a concentration of 0.15 M. The OPLS3e force field was employed to generate the MD system. MD simulations were performed using Desmond, with NPT ensembles at a temperature of 310 K and a pressure of 1.01325 bar. The system was relaxed using Desmond's default protocol before the MD production run. Each MD validation calculation was run for 2 ns.

4.6.8 Preparation of recombinant *Ng*CysE

*Ng*CysE was cloned and expressed using methods published previously (Oldham et al., 2022). *Ng*CysE was cloned into expression vector pET28b-PstI and expressed in *E. coli* BL21 (DE3) with a N-terminal HexaHistag. *Ng*CysE was cultured in 1L of LB at 37°C and induced with Isopropyl β- d-1-thiogalactopyranoside IPTG (0.75 mM final concentration) once an 0.4-0.6 OD A_{600} was reached. Cultures were left to express at 37°C overnight (180 rpm) and pellets were stored at -80°C before purification. Thawed pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 20 mM imidazole) with the addition of one cOmplete Mini EDTA free protease inhibitor tablet (Roche). On ice, the resuspended pellet was sonicated using a ¼ inch probe, in one second bursts with one second intervals, for 1.5 minutes of sonication in total, followed by centrifugation for 20 minutes at 13,945 xg (10°C). For IMAC purification, 0.22 µM filtered supernatant was loaded onto a 5 ml HisTrap^{TM}HP column (Cytiva) and eluted over a 50% imidazole gradient (50 mM Tris pH 8.0, 200 mM NaCl, 1 M imidazole) eluting at a final concentration of 400 mM imidazole. Eluted protein was concentrated at 3000 xg at 10°C using a 10K MWCO concentrator and loaded onto an analytical Enrich650 size exclusion column (Bio-Rad) pre-equilibrated in size exclusion buffer (50 mM Tris pH 8.0, 100 mM NaCl), where it eluted as a single peak with an elution volume of 12.9 ml consistent with a hexameric oligomer. Protein concentration was measured through measuring absorbance at A280nm ($\varepsilon = 0.598$) and diluted to an appropriate working

stock concentration for *Ng*CysE activity assays and inhibitor screening. All assays were conducted in size exclusion buffer.

4.6.9 *Ng*CysE activity assays and inhibitor screening

*Ng*CysE activity and inhibitor screening was measured in a 96-well plate format using a coupled DTNB assay monitoring the formation of CoA producing the coloured product TNB detected at 412 nm (ϵ = 14,150 M^{-1} cm^{-1}) (Riddles, Blakeley, & Zerner, 1983). All assays were collected in 50 mM Tris pH 8.0, 100 mM NaCl at 25°C with 1 mM DTNB, 0.1 mM acetyl-CoA, 1 mM serine and 1.97 nM of purified *Ng*CysE, in a final reaction volume of 250 µl, unless otherwise stated. For inhibitor screening, substrate concentrations were kept at or below K_m (serine K_m = 1.21 mM, acetyl-CoA K_m = 0.15 mM (Oldham et al., 2022)) to allow for detection of inhibitors with a competitive mode of inhibition. Reactions were measured for 20 minutes and the initial linear velocity was determined from the linear portion of the reaction.

Inhibition assays were optimised through pre-incubation of the enzyme with inhibitors before initiating the reaction with a substrate mix of serine and acetyl-CoA. All compound stocks were prepared in 100% DMSO with 1-10% DMSO in the final reaction volume. All positive controls contained an appropriate amount of DMSO in the final reaction. All experimentally tested compounds were ordered with at least 90% purity from commercial vendors ChemBridge, MolPort and Asinex. Compound 2 was purchased from ChemBridge (Catalogue number 11296740, ZINC15 number ZINC11882369). *In vitro* screening of compounds at 100 µM was collected in the presence of 0.15 mM acetyl-CoA, 1 mM serine. IC_{50} dose response curves were collected in the presence of 0.1 mM acetyl-CoA and 0.5 mM serine through fitting the data to the non-linear regression dose-response curve – variable slope (Equation 4.1) using Prism (GraphPad, version 10.1.0).

Equation 4.1 Inhibitor vs. response -Variable slope

$$V = \frac{Bottom + (Top - Bottom)}{((1 + \frac{IC_{50}}{[I]})^{hillslope})}$$

4.7 Acknowledgements

Research reported here was funded by the Maurice Wilkins Centre for Molecular Biodiscovery and a Health Research Council Emerging Researcher Grant to J.H. Investigation: K.E.O. and W.J.: K.E.O., W.J., and J.H; writing – original draft: writing – review and editing; supervision: J.H.; funding acquisition: J.H.

4.8 Supplementary Materials

Supplementary Tables and Figures present in this manuscript are located in Appendix B: Supplementary Material for Chapter Four.

4.9 Future directions

4.9.1 Further characterisation of *Ng*CysE inhibition by compound 2

This work demonstrates crucial progress into the development of novel inhibitors for serine acetyltransferase from *Neisseria gonorrhoeae* (*Ng*CysE). Although the inhibition we observe for our most potent inhibitor, compound 2, is lower than the affinity of native inhibitor L-cysteine (Oldham et al., 2022), the inhibition observed is within magnitude. The identification of compound 2 demonstrates that structure-based virtual inhibitor screening is a suitable approach for identifying *Ng*CysE inhibitors. This work lays the foundations for and strongly supports future work characterising inhibitory nature of compound 2 to improve the potency of this promising inhibitor.

In this work, we identified compound 2 to exhibit inhibition in the low micromolar range but ideally for antimicrobial candidate, inhibition in the nanomolar range is ideal and so hit optimisation is needed. As a part of future research, we are going to conduct medicinal chemistry screening of compound 2 with collaborators, Dr. Farah Lamiable-Oulaidi and Dr. Wanting Jiao, at the Ferrier Research Institute at Victoria the University of Wellington. This work would allow us to conduct structure-activity relationship studies of this compound, to determine which functional groups and residues interactions are needed for potent inhibition.

To compliment the structure-activity relationship studies we will further characterise the inhibition of compound 2, including mode of inhibition and binding kinetics. Additionally, to improve the quality of the medicinal chemistry

screen, co-crystallisation of compound 2 with *Ng*CysE will be carried out to experimentally validate compound-*Ng*CysE interactions. Progress has already been made in this area, as we sought to identify a new crystallisation condition to obtain an apoenzyme *Ng*CysE structure without the presence of L-malate to increase the likelihood of compound 2 co-crystallising in the active site with full occupancy. Using robot screens we have identified and optimised a new crystallisation condition (without L-malate) for *Ng*CysE and have already collected, processed and analysed a new crystal *Ng*CysE structure (Figure 4.9) with comparable resolution (2.05 Å) to our original L-malate structure used for virtual inhibitor screening (6WYE, 2.01 Å). This new crystallisation condition will be suitable for co-crystallisation/soaking experiments for compound 2 and will allow us to confirm the interaction we have seen in virtual inhibitor docking, which in turn will help inform our medicinal chemistry efforts to screen compound functional groups and analogues and determine the structure activity relationship for this compound.

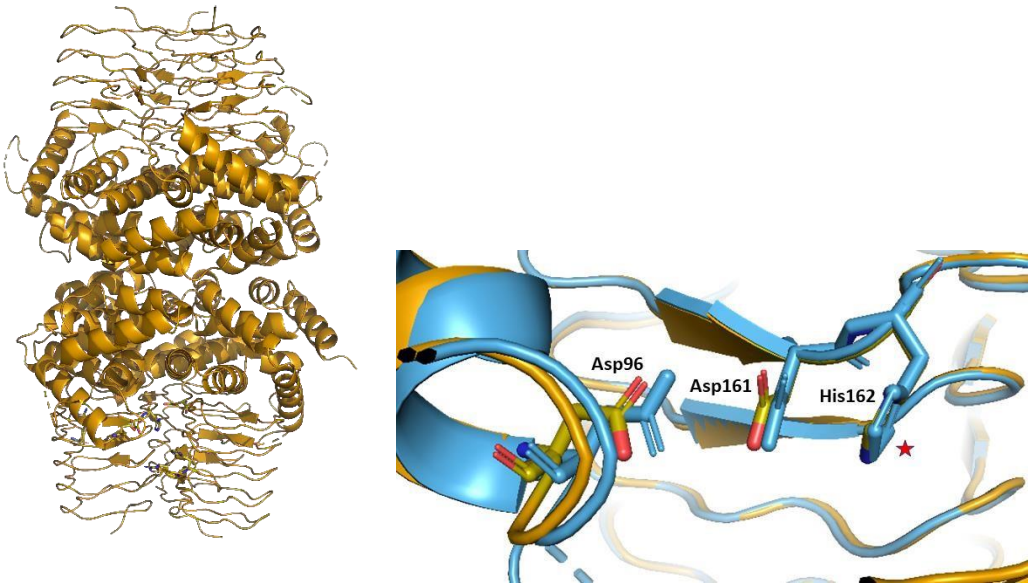


Figure 4.9 Alignment of new apoenzyme *Ng*CysE model to *Ng*CysE malate structure (6WYE). *Ng*CysE crystallised as a homohexamer with chains displayed as cartoon. Alignment of apoenzyme *Ng*CysE (orange) to *Ng*CysE malate (blue) (RMSD 0.534). Active site residues are shown as sticks. Waters and malate are hidden for clarity.

Given the strong conservation of serine and acetyl-CoA binding site residues amongst CysE bacterial homologues, particularly amongst Proteobacteria members (Figure 4.7). Cross species inhibition between highly similar isoforms is

possible and the screening *in vitro* against one isoform and testing for bioactivity against a different species has been used to successfully identify CysE inhibitors (Toyomoto et al., 2023). Additionally, in the absence of a crystal structure of other target species, structures from closely related isoforms have been used for virtual screening to identify compounds with *in vitro* inhibition (Magalhães et al., 2020). Given that we have established a protocol for expression, purification, and characterisation of CysE enzymes in our lab, we are well-suited for characterisation and inhibitor screening of other bacterial CysE isoforms. To identify new inhibitory compounds, the screening workflow that we have established in our lab and with collaborators, can be used for future virtual and *in vitro* inhibitor screening. In addition, *in vitro* screening of our hit compound libraries reported in this thesis chapter could be screened against other CysE isoforms, particularly those with high degree of similarity. As mentioned above in the manuscript discussion, given the high degree of sequence similarity between *Nm*CysE and *Ng*CysE this would be a well-suited candidate and would be a natural extension of this work. Furthermore, considering the differences in the sulphur assimilation pathways between *N. gonorrhoeae* and *N. meningitidis* (Hicks & Mullholland, 2018), the biochemical and *in vivo* characterisation of *Nm*CysE, particularly given the essential nature of *cysE* in *N. meningitidis*, would help elucidate the role of *de novo* cysteine in pathogenic *Neisseria*.

4.9.2 Characterisation of *Neisseria gonorrhoeae cysE* knockout strain

In vivo characterisation of an antimicrobial target is essential for drug target validation and success. CysE is encoded by the *cysE* gene in *N. gonorrhoeae* and was identified as an essential gene in *N. gonorrhoeae* MS11 by a transposon mutagenesis study (Remmele et al., 2014). Assuming the essential nature of *cysE* to be correct, we set out to create a *cysE* depleted strain for *cysE* characterisation. We initially designed a construct where the native *cysE* promoter would be swapped with a tetracycline-inducible promoter along with the tetracycline repressor and a kanamycin selection marker (taken from plasmid pMR68, (Ramsey, Hackett, Kotha, & Dillard, 2012). Given the natural competency of *N. gonorrhoeae*, piliated *N. gonorrhoeae* was streaked through spotted DNA on gonococcal broth (GCB) agar supplemented with Isovitalex (1% final volume, contains L-cysteine and L-cystine) and anhydrotetracycline (20 $ng.ml^{-1}$) following an adapted spot

transformation protocol (Dillard, 2011). However, the construct only yielded mixed colonies (colony PCR products sizes indicating presence of both WT and integration of construct) or given the polyploid nature of the *N. gonorrhoeae* genome (Tobiason & Seifert, 2006) the integration might have occurred in only some chromosome copies. Furthermore, these colonies were not amenable to freeze-thawing and were not revivable from glycerol stocks and so efforts were made to find alternative strategies for generating a *cysE* depleted strain.

Our second attempt involved creating a *cysE* depleted strain through first introducing a copy of *cysE* under the control of a inducible promoter (lac promoter, from pKH37 plasmid) with a resistance marker (CAT, chloramphenicol) at a neutral site in the genome (intergenic region between *iga* (immunoglobulin A protease) and *trpB* (tryptophan synthase beta chain) used for complementation (Ramsey et al., 2012) and then subsequently knockout *cysE* from the native site with a different resistance marker (KanR, kanamycin). We chose to complement by inserting *cysE* into the chromosome as replicating plasmids are unstable for *N. gonorrhoeae* (Dillard, 2011). *cysE* was amplified from gDNA isolated from *N. gonorrhoeae* MS11 and cloned into the pKH37 plasmid for spot transformation. This method proved successful, and we were able to delete the genomic copy of *cysE* from the *N. gonorrhoeae* MS11 inducible *cysE* strain (Ramsey et al., 2012). Additionally, our attempts to knockout *cysE* from wildtype *N. gonorrhoeae* MS11 were unsuccessful and correspondence with Dr. Linda Hu based in Professor Hank Seifert's lab at Northwestern University reported they were unable to knockout *cysE* from the *N. gonorrhoeae* FA1090 strain. Interestingly, we report that this *cysE* inducible knockout can grow in the absence of *cysE* induction by IPTG on GCB agar. This could be due to one of two reasons; (1) *cysE* is conditionally essential as GCB is a rich media and will contain the amino acids, cysteine and cystine, which can be actively imported into the cell and is able to meet cysteine requirements (Bulut et al., 2012) or (2) due to the leaky nature of the *lac* promoter in the absence of IPTG there is *cysE* expression which is able to meet the CysE levels needed for *N. gonorrhoeae* growth. Overall, given our inability to knockout *cysE* from WT *N. gonorrhoeae* MS11, future work will involve creating a new *cysE* depletion strain. This will involve the same strategy of pre-complimenting at the same neutral site but will use a tetracycline inducible promoter and repressor

(from the same plasmid previously used for promoter swap, pMR68) to reduce the amount of leaky expression and will measure *cysE* expression through CysE levels using western blots. Overall, this is an excellent starting point for *in vivo* characterisation of *Ng*CysE and a *cysE* depleted strain will be crucial for understanding the role of *cysE* in *N. gonorrhoeae*.

4.10 References

Archibald, F. S., & Duong, M. N. (1986). Superoxide dismutase and oxygen toxicity defenses in the genus Neisseria. *Infection and immunity, 51*(2), 631-641.

Baell, J. B., & Holloway, G. A. (2010). 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*(7), 2719-2740. 10.1021/jm901137j

Bajusz, D., Rácz, A., & Héberger, K. (2015). Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J Cheminform, 7*, 20. 10.1186/s13321-015-0069-3

Benoni, R., De Bei, O., Paredi, G., Hayes, C. S., Franko, N., Mozzarelli, A., . . . Campanini, B. (2017). Modulation of Escherichia coli serine acetyltransferase catalytic activity in the cysteine synthase complex. *FEBS letters, 591*(9), 1212-1224. 10.1002/18733468.12630

Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., . . . Shaw, D. E. (2006). Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing* (pp. 43-43): 10.1109/SC.2006.54

Bulut, H., Moniot, S., Licht, A., Scheffel, F., Gathmann, S., Saenger, W., & Schneider, E. (2012). Crystal Structures of Two Solute Receptors for l-Cystine and l-Cysteine, Respectively, of the Human Pathogen Neisseria gonorrhoeae. *Journal of Molecular Biology, 415*(3), 560-572. 10.1016/j.jmb.2011.11.030

Campanini, B., Pieroni, M., Raboni, S., Bettati, S., Benoni, R., Pecchini, C., . . . Mozzarelli, A. (2015). Inhibitors of the Sulphur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Current Medicinal Chemistry, 22*(2), 187-213. 10.2174/0929867321666141112122553

Capel, E., Zomer, A. L., Nussbaumer, T., Bole, C., Izac, B., Frapy, E., . . . Coureuil, M. (2016). Comprehensive Identification of Meningococcal Genes and Small Noncoding RNAs Required for Host Cell Colonization. *mBio, 7*(4), e01173-01116. 10.1128/mBio.01173-16

Chen, C., Yan, Q., Tao, M., Shi, H., Han, X., Jia, L., . . . Ma, Y. (2019). Characterization of serine acetyltransferase (CysE) from methicillin-resistant Staphylococcus aureus and inhibitory effect of two natural products on CysE. *Microbial Pathogenesis, 131*, 218-226. 10.1016/j.micpath.2019.04.002

Dietrich, G., Kurz, S., Hübner, C., Aepinus, C., Theiss, S., Guckenberger, M., . . . Frosch, M. (2003). Transcriptome analysis of Neisseria meningitidis during infection. *J Bacteriol, 185*(1), 155-164. 10.1128/jb.185.1.155-164.2003

Dillard, J. P. (2011). Genetic Manipulation of Neisseria gonorrhoeae. *Current protocols in microbiology, Chapter 4*, Unit4A.2-Unit4A.2. 10.1002/9780471729259.mc04a02s23

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 4), 486-501. 10.1107/s0907444910007493

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., . . . Shenkin, P. S. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry, 47*(7), 1739-1749. 10.1021/jm0306430

Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., . . . Mainz, D. T. (2006). Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *Journal of Medicinal Chemistry, 49*(21), 6177-6196. 10.1021/jm051256o

Greenwood, J. R., Calkins, D., Sullivan, A. P., & Shelley, J. C. (2010). Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of Computer-Aided Molecular Design, 24*(6), 591-604. 10.1007/s10822-010-9349-1

Grifantini, R., Bartolini, E., Muzzi, A., Draghi, M., Frigimelica, E., Berger, J., . . . Grandi, G. (2002). Gene Expression Profile in Neisseria meningitidis and Neisseria lactamica upon Host-Cell Contact. *Annals of the New York Academy of Sciences, 975*(1), 202-216. 10.1111/j.1749-6632.2002.tb05953.x

Griffith, O. W. (1987). Mammalian sulphur amino acid metabolism: An overview *Methods in Enzymology* (Vol. 143, pp. 366-376): Academic Press.

Gupta, S., & Gupta, V. (2021). Homology modeling, structural insights and in-silico screening for selective inhibitors of mycobacterial CysE. *J Biomol Struct Dyn, 39*(5), 1547-1560. 10.1080/07391102.2020.1734089

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry, 47*(7), 1750-1759. 10.1021/jm030644s

Hicks, J. L., & Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology, 164*(12), 1471-1480. 10.1099/mic.0.000728

Hicks, J. L., Oldham, K. E. A., McGarvie, J., & Walker, E. J. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. *Biosci Rep, 42*(10) 10.1042/bsr20220368

Jacobson, M. P., Friesner, R. A., Xiang, Z., & Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol, 320*(3), 597-608. 10.1016/s0022-2836(02)00470-9

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J., Honig, B., Shaw, D. E., & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins, 55*(2), 351-367. 10.1002/prot.10613

Jolley, K. A., Bray, J. E., & Maiden, M. C. J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res, 3*, 124. 10.12688/wellcomeopenres.14826.1

Joseph, B., Schneiker-Bekel, S., Schramm-Glück, A., Blom, J., Claus, H., Linke, B., . . . Schoen, C. (2010). Comparative genome biology of a serogroup B carriage and disease strain supports a polygenic nature of meningococcal virulence. *J Bacteriol, 192*(20), 5363-5377. 10.1128/jb.00883-10

Kredich, N. (2008). Biosynthesis of Cysteine. *EcoSal Plus* doi:10.1128/ecosalplus.3.6.1.11

Kredich, N. M., Becker, M. A., & Tomkins, G. M. (1969). Purification and Characterization of Cysteine Synthetase, a Bifunctional Protein Complex, from Salmonella typhimurium. *Journal of Biological Chemistry, 244*(9), 2428-2439.

Kredich, N. M., & Tomkins, G. M. (1966). The Enzymic Synthesis of l-Cysteine in Escherichia coli and Salmonella typhimurium. *Journal of Biological Chemistry, 241*(21), 4955-4965.

Landrum, G. (2023). RDKit: Open-source cheminformatics (Version Release_2023.09.3). Retrieved from https://www.rdkit.org

Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *Journal of Chemical Information and Modeling, 51*(10), 2778-2786. 10.1021/ci200227u

Li, H., Robertson, A. D., & Jensen, J. H. (2005). Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Structure, Function, and Bioinformatics, 61*(4), 704-721. 10.1002/prot.20660

Lipinski, C. A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods, 44*(1), 235-249. 10.1016/s10568719(00)00107-6

Magalhães, J., Franko, N., Raboni, S., Annunziato, G., Tammela, P., Bruno, A., . . . Costantino, G. (2021). Discovery of Substituted (2-Aminooxazol-4-yl)Isoxazole3-carboxylic Acids as Inhibitors of Bacterial Serine Acetyltransferase in the Quest for Novel Potential Antibacterial Adjuvants. *Pharmaceuticals, 14*(2), 174.

Magalhães, J., Franko, N., Raboni, S., Annunziato, G., Tammela, P., Bruno, A., . . . Costantino, G. (2020). Inhibition of Nonessential Bacterial Targets: Discovery of a Novel Serine O-Acetyltransferase Inhibitor. *ACS Medicinal Chemistry Letters* 10.1021/acsmedchemlett.9b00627

Maher, C., & Hassan, K. A. (2023). The Gram-negative permeability barrier: tipping the balance of the in and the out. *mBio, 14*(6), e01205-01223. doi:10.1128/mbio.01205-23

Muir, A., Gurung, I., Cehovin, A., Bazin, A., Vallenet, D., & Pelicic, V. (2020). Construction of a complete set of Neisseria meningitidis mutants and its use for the phenotypic profiling of this human pathogen. *Nature Communications, 11*(1), 5541. 10.1038/s41467-020-19347-y

Murphy, R. B., Philipp, D. M., & Friesner, R. A. (2000). A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *Journal of Computational Chemistry, 21*(16), 1442-1457. 10.1002/1096-987X(200012)21:16<1442::AID-JCC3>3.0.CO;2-O

Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., . . . Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet, 399*(10325), 629-655. 10.1016/S01406736(21)02724-0

Oldham, K. E. A., Prentice, E. J., Summers, E. L., & Hicks, J. L. (2022). Serine acetyltransferase from Neisseria gonorrhoeae; structural and biochemical basis of inhibition. *Biochem J, 479*(1), 57-74. 10.1042/bcj20210564

Park, S., & Imlay, J. A. (2003). High Levels of Intracellular Cysteine Promote Oxidative DNA Damage by Driving the Fenton Reaction. *Journal of Bacteriology, 185*(6), 1942-1950. 10.1128/jb.185.6.1942-1950.2003

Ramsey, M. E., Hackett, K. T., Kotha, C., & Dillard, J. P. (2012). New Complementation Constructs for Inducible and Constitutive Gene Expression in Neisseria gonorrhoeae and Neisseria meningitidis. *Applied and Environmental Microbiology, 78*(9), 3068-3078. 10.1128/aem.07871-11

Remmele, C. W., Xian, Y., Albrecht, M., Faulstich, M., Fraunholz, M., Heinrichs, E., . . . Rudel, T. (2014). Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic acids research, 42*(16), 10579-10595. 10.1093/nar/gku762

Riddles, P. W., Blakeley, R. L., & Zerner, B. (1983). [8] Reassessment of Ellman's reagent *Methods in Enzymology* (Vol. 91, pp. 49-60): Academic Press.

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des, 27*(3), 221-234. 10.1007/s10822-0139644-8

Schrödinger. (2020a). Desmond Molecular Dynamics System. New York, NY: D. E. Shaw Research.

Schrödinger. (2020b). Epik (Version 2020-02). Retrieved from http://www.schrodinger.com/

Schrödinger. (2020c). Glide (Version 2020-02). New York, NY: Schrödinger.

Schrödinger. (2020d). LigPrep. New York, NY: Schrödinger.

Schrödinger. (2020e). Maestro (Version Release 2020-02): Schrödinger. Retrieved from http://www.schrodinger.com/

Schrödinger. (2020f). Prime (Version Release 2020-02): Schrödinger. Retrieved from http://www.schrodinger.com/

Schrödinger. (2020g). Protein Preparation Wizard (Version 2020-02). New York, NY: Schrödinger.

Schrödinger. (2020h). QSite (Version 2020-02). New York, NY.

Seib, K. L., Wu, H.-J., Kidd, S. P., Apicella, M. A., Jennings, M. P., & McEwan, A. G. (2006). Defenses against oxidative stress in Neisseria gonorrhoeae: a system tailored for a challenging environment. *Microbiology and molecular biology reviews : MMBR, 70*(2), 344-361. 10.1128/MMBR.00044-05

Semchenko, E. A., Day, C. J., & Seib, K. L. (2017). MetQ of Neisseria gonorrhoeae Is a Surface-Expressed Antigen That Elicits Bactericidal and Functional Blocking Antibodies. *Infection and Immunity, 85*(2), 10.1128/iai.00898-00816. doi:10.1128/iai.00898-16

Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: a software program for pKaprediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design, 21*(12), 681-691. 10.1007/s10822-007-9133-z

Sterling, T., & Irwin, J. J. (2015). ZINC 15--Ligand Discovery for Everyone. *J Chem Inf Model, 55*(11), 2324-2337. 10.1021/acs.jcim.5b00559

Takahashi, H., & Watanabe, H. (2005). A gonococcal homologue of meningococcal gamma-glutamyl transpeptidase gene is a new type of bacterial pseudogene that is transcriptionally active but phenotypically silent. *BMC Microbiol, 5*, 56. 10.1186/1471-2180-5-56

Talà, A., Calcagnile, M., Resta, S. C., Pennetta, A., De Benedetto, G. E., & Alifano, P. (2023). Thiostrepton, a resurging drug inhibiting the stringent response to counteract antibiotic-resistance and expression of virulence determinants in Neisseria gonorrhoeae. *Frontiers in Microbiology, 14* 10.3389/fmicb.2023.1104454

Tobiason, D. M., & Seifert, H. S. (2006). The Obligate Human Pathogen, Neisseria gonorrhoeae, Is Polyploid. *PLOS Biology, 4*(6), e185. 10.1371/journal.pbio.0040185

Toyomoto, T., Ono, K., Shiba, T., Momitani, K., Zhang, T., Tsutsuki, H., . . . Sawa, T. (2023). Alkyl gallates inhibit serine O-acetyltransferase in bacteria and enhance susceptibility of drug-resistant Gram-negative bacteria to antibiotics. *Frontiers in Microbiology, 14* 10.3389/fmicb.2023.1276447

Unemo, M., Golparian, D., & Hellmark, B. (2014). First Three Neisseria gonorrhoeae Isolates with High-Level Resistance to Azithromycin in Sweden: a Threat to Currently Available Dual-Antimicrobial Regimens for Treatment of Gonorrhea? *Antimicrobial Agents and Chemotherapy, 58*(1), 624-625. doi:10.1128/aac.0209313

Verma, D., Gupta, S., Saxena, R., Kaur, P., R, R., Srivastava, S., & Gupta, V. (2020). Allosteric inhibition and kinetic characterization of Klebsiella pneumoniae CysE: An emerging drug target. *Int J Biol Macromol, 151*, 1240-1249. 10.1016/j.ijbiomac.2019.10.170

WHO. (2017). World Health Organization (WHO) Global Priority List of Antibiotic Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics. Geneva: WHO.

Wi, T., Lahra, M. M., Ndowa, F., Bala, M., Dillon, J.-A. R., Ramon-Pardo, P., . . . Unemo, M. (2017). Antimicrobial resistance in Neisseria gonorrhoeae: Global surveillance and a call for international collaborative action. *PLOS Medicine, 14*(7), e1002344. 10.1371/journal.pmed.1002344

Chapter Five: Distribution and diversity of serine acetyltransferases from bacteria

5.1 Preface

This chapter reports our investigations into the phylogenetic distribution of truncated serine acetyltransferases (CysE) in bacteria. We demonstrate that the truncated CysE isoform has a consistent N-terminal truncation of the first four αhelices of the N-terminal serine acetyltransferase domain (IPR042122). Further, we show the truncated isoform is the dominant isoform for Gram-positive bacterial species, while the full-length CysE dominates Proteobacterial members. Lastly, using sequence similarity networks we report individual clustering of CysE sequence into full-length and truncated isoforms. Finally, we analyse the implications of the truncation on the CysE oligomeric state through predictive modelling.

This work is prepared for submission as a Short Communication to the journal *PROTEINS: Structure, Function and Bioinformatics* (Wiley). Included in this chapter is a Future Directions section (5.8), which itself is not part of the manuscript, but has been included to discuss ongoing and future work directly related to this project; these results will be included in the future publications. Supplementary material associated with this manuscript is available in Appendix C. Citation for this work is as follows:

Oldham, K. E. A., Grout, E. K., Williamson, A., & Hicks, J. L. The curious case of CysE: diversity and distribution of serine acetyltransferases. *(Prepared for submission)*

5.1.1 Author Contributions

As first author, I was responsible for setting research questions and the scope of this article. Further, I conducted the bioinformatic work, predictive modelling of the CysE isoforms, figure preparation, as well as manuscript drafting and editing. Emily Grout was responsible for phylogenetic analysis and figure preparation. Adele Williamson performed analysis of the phylogenetics and manuscript editing. Joanna Hicks was responsible for setting the scope of this article, project

supervision, funding acquisition, as well as manuscript drafting and editing. The associated co-authorship form can be found in Appendix D.

The curious case of CysE: diversity and distribution of serine acetyltransferases

Keely E.A. Oldham[1,2], Emily K. Grout[1], Adele Williamson[2], Joanna L. Hicks[1]

[1]*Te Huataki Waiora, School of Health, University of Waikato, Hamilton, New Zealand*

[2]*Te Aka Mātuatua, School of Science, University of Waikato, Hamilton, New Zealand*

KEYWORDS

Serine acetyltransferase, CysE, SAT, truncation, hexapeptide repeat, cysteine biosynthesis

5.2 Abstract

Serine acetyltransferase (CysE) is a member of the left-handed β-helix family of acetyltransferases that catalyses the rate limiting step in the *de novo* cysteine biosynthesis pathway. There are two isoforms of CysE, one slightly longer than the other, with the shorter, truncated isoform lacking approximately 76 amino acids at the N-terminus of the protein. Here, we analyse the distribution and diversity of CysE isoforms across the bacterial kingdom. The isoforms can be classified into two discrete groups, with the truncated isoform prevalent in Gram-positive bacteria and the full-length isoform prevalent in Gram-negative species. Moreover, we demonstrate that the truncation is discrete with the loss of four N-terminal alpha helices conserved amongst this isoform. Using predictive modelling we demonstrate that this truncation likely weakens the CysE hexamer, leading to a trimeric form instead of the canonical CysE hexamer. This expands our understanding of CysE enzymes and their distribution across bacterial species, an important consideration given the increasing interest in targeting CysE enzymes for potential antimicrobials.

5.3 Introduction

The *de novo* biosynthesis of cysteine is conserved in bacteria, higher plants and protozoa. Cysteine is a key amino acid and important for a variety of functions

within the cell. For example, L-cysteine feeds into the synthesis of co-enzyme A, iron-sulphur clusters, biotin, and reducing agents such as thioredoxin and glutathione (Kredich, 2008). Accordingly, cysteine is especially important for bacterial pathogens to mitigate oxidative stress encountered during infection. The biosynthesis of L-cysteine begins with the active transport of a sulphur donor, either sulphate or thiosulphate, into the cell through an ABC-transporter. Sulphate is successively reduced to sulphide which is incorporated into cysteine whereas thiosulphate is directly used for the synthesis of cysteine. The two-step cysteine biosynthetic pathway begins with the enzyme, serine acetyltransferase (SAT; CysE, EC.3.2.10) which catalyses the acetylation of L-serine, producing pathway intermediate *O*-acetylserine. In the second step, *O*-acetylserine either undergoes a condensation reaction with sulphide catalysed by *O*-acetylserine sulphydrylase-A (OASS-A; CysK) or thiosulphate catalysed by *O*-acetylserine sulphydrylase-B (OASS-B; CysM) to form L-cysteine. Cellular cysteine levels are tightly controlled, to avoid toxicity and prevent the unwanted Fenton reaction (Park & Imlay, 2003). CysE is inhibited by classical feedback inhibition by L-cysteine (Hindson, 2003), and the two enzymes in the pathway, CysE and CysK, form the cysteine synthase complex (CSC) consisting of one CysE hexamer and two CysK dimers (Benoni et al., 2017). In the complex, CysK activity is inhibited preventing the synthesis of L-cysteine, whereas CysE activity is enhanced and leads to increasing levels of pathway intermediate *O*-acetylserine (Benoni et al., 2017). High concentrations of *O*-acetylserine disrupt the complex, restoring CysK activity resulting in synthesis of cysteine. Additionally, the complex is stabilised by sulphide. Collectively the complex acts to regulate sulphur flux within the cell (Benoni et al., 2017).

CysE is a member of the hexapeptide acyltransferase family. The CysE monomer consists of two domains; an alpha-helical N-terminal serine acetyltransferase domain (IPR042122, serine acetyltransferase N-terminal domain superfamily) which is unique to CysE enzymes and a C-terminal left-handed β-helix (LβH) domain (IPR011004, trimeric LpxA-like superfamily) formed through the presence of a conserved hexapeptide repeat [LIV]-[GAED]-X_2-[STAV]-X and is shared with other members of the LpxA acetyltransferases (Raetz & Roderick, 1995; Vaara, 1992). The monomers assemble to form a hexameric dimer of trimers, observed both in crystal structures (Gorman & Shapiro, 2004; Oldham, Prentice, Summers,

& Hicks, 2022; Olsen, B., Vetting, & Roderick, 2004; Toyomoto et al., 2023; Verma et al., 2020) and in solution (Benoni et al., 2017; Oldham et al., 2022). This hexameric assembly is unique to CysE, with other members of the LβH acyltransferase family adopting a trimeric configuration (Raetz & Roderick, 1995). The CysE active site is formed within an inter-monomer cleft within the trimer and consists of highly conserved serine and acetyl-CoA binding residues. The serine binding pocket contains a catalytic dyad consisting of a histidine and an aspartate (His158 and Asp92, *E. coli* numbering) that together catalyse the transfer of the acetyl group from acetyl-CoA to the serine substrate via a random ternary complex mechanism (Hindson & Shaw, 2003; Pye, Tingey, Robson, & Moody, 2004).

Recently, research characterising bacterial CysE homologues has uncovered the presence of truncated isoforms (Chen et al., 2019; Qiu, Ma, Owusu, Jiang, & Xin, 2014; Qiu, Wang, Ma, Jiang, & Xin, 2013; Rahisuddin et al., 2024). While truncation is often associated with loss of catalysis, reported characterisation shows that the truncated isoforms retain their serine acetyltransferase activity although with lower rates (Table C.1). While the truncation has been previously identified (Gagnon et al., 1994; Gorman & Shapiro, 2004; Rahisuddin et al., 2024), it is unclear how widely distributed this truncated isoform is across bacteria. Furthermore, given the importance of N-terminal SATase domain for hexamer assembly (Gorman & Shapiro, 2004; Pye et al., 2004) the truncation is likely to impact the oligomeric species that CysE can adopt. The growing interest in targeting the *de novo* cysteine biosynthesis pathway for antimicrobial development (Hicks, Oldham, McGarvie, & Walker, 2022) has highlighted the understudied diversity of CysE isoforms, particularly in bacterial pathogens.

Here, we report the distribution and diversity of CysE isoforms across the bacterial kingdom. We identify the presence of two distinct CysE isoforms, defined by the length of the α-helical N-terminal SATase domain. We demonstrate that the isoform with a truncated α-helical region is the predominant isoform for Gram-positive bacteria and is largely absent from Gram-negative bacteria and explore the relationship between truncated and full-length isoforms using sequence similarity networks (SSNs). Lastly, in the absence of structures of the truncated isoform, we have used the predictive modelling tool ColabFold to examine the effect of the N-terminal truncation and its direct impacts on the multimeric assembly of CysE.

5.4 Results and Discussion

5.4.1 Analysis of the truncated CysE isoform

To examine the CysE isoforms we created a multiple sequence alignment (MSA) of truncated and full-length bacterial CysE isoforms. Sequences for full-length isoforms from crystal structures within the protein data bank (PDB) (*Escherichia coli* (*Ec*CysE), *Brucella abortus* (*Ba*CysE), *Salmonella typhimurium* (*St*CysE), *Haemophilus influenzae* (*Hi*CysE), *Neisseria gonorrhoeae* (*Ng*CysE), *Vibrio cholerae* (*Vc*CysE) and *Yersinia pestis* (*Yp*CysE)) and characterised truncated CysE isoforms (*Staphylococcus aureus* (*Sa*CysE) (Chen et al., 2019) and *Mycobacterium tuberculosis* (*Mtb*CysE) (Qiu et al., 2013)) were retrieved from UniProt. Sequence comparison of the truncated CysE isoforms to full-length sequences, revealed a good alignment of the C-terminal LβH domain (IPR045304) including hexapeptide repeats and active site residues (Figure 5.1). Whereas alignment of the N-terminal SATase domain (IPR042122) demonstrated the absence of the first four α-helices (α1-α4) of the N-terminal domain in *Sa*CysE and *Mtb*CysE isoforms CysE (Figure 5.1 and Figure 5.2) shortening the length of the N-terminal SATase domain from eight (~142 residues) to four α-helices (~66 residues) resulting in a loss of ~76 residues for these truncated isoforms.

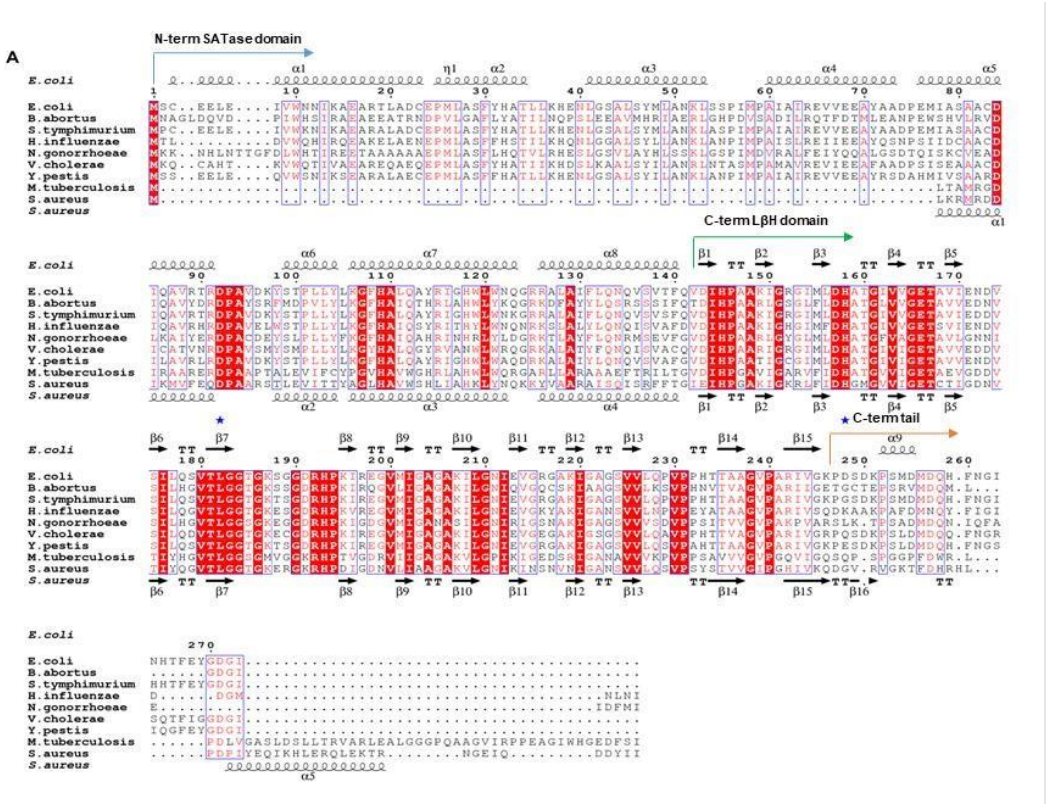


Figure 5.1 Structural comparison and MSA of full-length and truncated CysE isoforms. (A) MSA of characterised CysE bacterial isoforms. Secondary structure annotations above and below MSA, correspond to representative full-length isoforms of CysE from *E. coli* (*Ec*CysE; PDB 1T3D) and representative truncated CysE from *S. aureus* (*Sa*CysE; AlphaFold model AF-P6766-F1, no experimental structure available), respectively. Strictly conserved residues are highlighted in red, partially conserved residues are highlighted in red text and repeats are highlighted with blue boxes. Catalytic dyad residues Asp92 and His158 (*Ec*CysE residue numbering) are annotated with blue stars. Species used in MSA; *E. coli* (P0A9D4, PDB 1T3D), *Brucella abortus* (A0A0F6AR69, PDB 4HZC), *Salmonella typhimurium* (A0A0D6I3Y9, PDB 7E3Y), *Haemophilus influenzae* (P43886, PDB 1S80), *Neisseria gonorrhoeae* (Q5FX0, PDB 6WYE), *Vibrio cholerae* (Q9KNT2, PDB 4H7O), *Yersinia pestis* (A0A2U2H3H7, PDB 3GVD), *Mycobacterium tuberculosis* (P95231, no structure available) and *S. aureus* (P67766, AlphaFold model AF-P6766-F1). MSA created using ESPript 3.0 (Robert & Gouet, 2014).

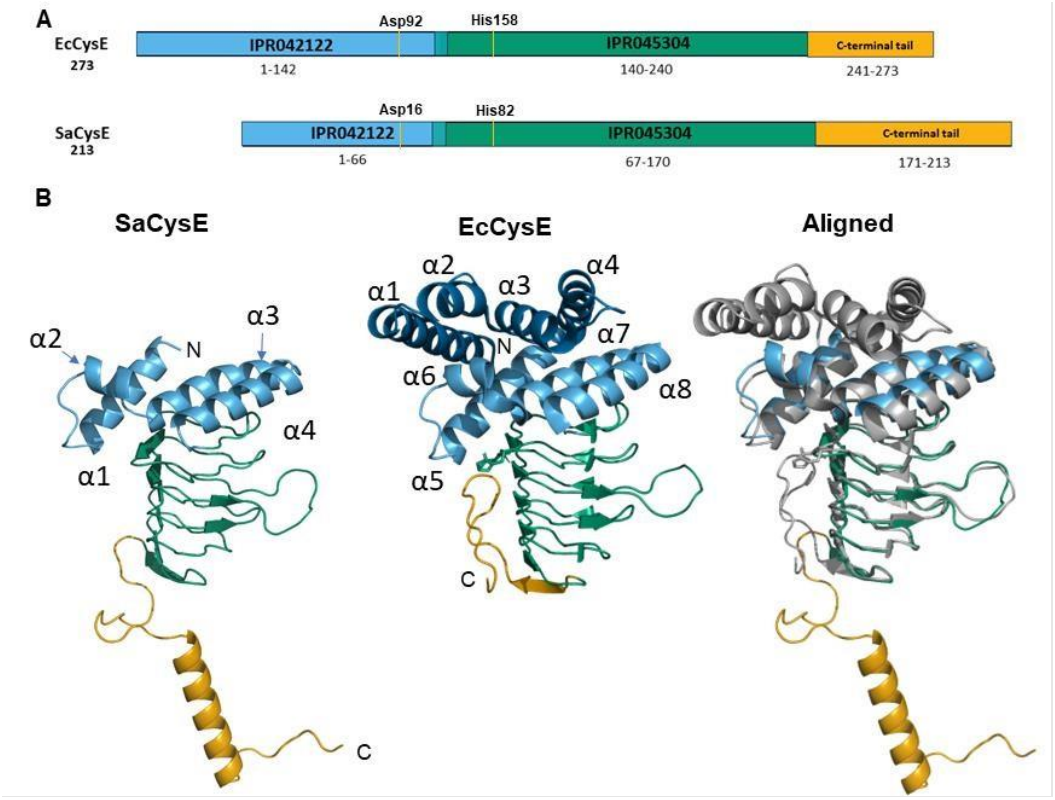


Figure 5.2 Structural alignment of truncated and full-length CysE isoforms. (A) Schematic of domains for *Ec*CysE and *Sa*CysE homologues. N-terminal SATase domain (IPR042122) is coloured blue, left-handed β-helix (IPR045304) is coloured green, and the C-terminal α-helical tail is coloured orange. Residue domain ranges are noted below respective domains. Catalytic dyad residues are indicated with yellow lines. Domain lengths are not drawn to scale. (B) AlphaFold-predicted model of the *Sa*CysE monomer (coloured based on scheme in A) and the crystal structure of *Ec*CysE (PDB ID 1T3D) indicated. *Ec*CysE α-helices (α1-α4) are shaded in dark blue for clarity. Structural overlay performed in PyMOL of the *Sa*CysE (coloured) and *Ec*CysE (grey) monomers highlighting the differences between isoforms.

In the absence of available crystal structures for these truncated isoforms, the predicted AlphaFold model was retrieved for the *Sa*CysE monomer from the AlphaFold database. The *Sa*CysE monomer was predicted with high confidence (Figure C.1) and structural alignment of *Sa*CysE to the full-length isoform *Ec*CysE showed good alignment of the LβH domain and SATase domain (RMSD 0.622), with the exception of the missing four α-helices, indicating structural conservation of the aligned regions (Figure 5.2). Additionally, both the MSA and AlphaFold model identified an elongated α-helix in the C-terminal tail in the truncated isoforms (*Sa*CysE and *Mtb*CysE) compared to the full-length isoforms (Figure 5.1 and Figure 5.2). The C-terminal tail is well-documented to contribute to both

feedback inhibition by cysteine (Mino, Hiraoka, et al., 2000; Pye et al., 2004) and formation of the cysteine synthase complex (Mino, Hiraoka, et al., 2000). However, there are no published attempts at measuring cysteine inhibition for truncated isoforms (Chen et al., 2019; Qiu et al., 2014; Qiu et al., 2013; Rahisuddin et al., 2024). Interestingly, the C-terminal isoleucine critical for forming the cysteine synthase complex is present in both *Sa*CysE and *Mtb*CysE but the effect of extended C-terminal tail on complex formation has not been explored.

5.4.2 CysE truncated isoforms cluster independently from full-length isoforms

The taxonomic distribution of CysE sequences reflects the conservation of *de novo* cysteine biosynthesis in bacteria. There are 39,897 sequences (21, 995 species) that are annotated with the SATase domain (IPR042122) in the UniProt database (Figure C.2). Sequences belonging to Archaea and Eukaryota were identified, but the majority of sequences are bacterial (Figure C.2). The taxonomic distribution of CysE reflects the observation that *de novo* cysteine biosynthesis is well-conserved across a wide range of bacterial phyla (Kredich, 2008). The *de novo* cysteine biosynthesis pathway is absent from the majority of eukaryotes with the exception of higher plants (Yi et al., 2013) and protozoa (Kumar, Raj, Nagpal, Subbarao, & Gourinath, 2011; Nozaki et al., 1999). Although, archaeal species have been reported to use an alternate tRNA-dependent SepRs/SepCysS pathway for cysteine biosynthesis (Sauerwald et al., 2005), some archaeal methanogens have been shown to contain CysE and CysK isoforms (Kitabatake, So, Tumbula, & Söll, 2000) (Figure C.2).

To determine the similarity between truncated and full-length CysE isoforms, SSNs were generated using the EFI-EST server. Sequences containing the N-terminal SATase domain (IPR042122) domain were used to construct the SSN. Initial analysis of the sequences revealed that the length of the SATase domain had a bimodal distribution, with the majority of sequences having a domain length between 52-76 or 130-166 residues (Figure C.3). This distribution is consistent with the observed domains sizes for truncated and full-length isoforms respectively (network details are summarised in Table C.2).

The initial SSN contained 5095 repnodes (1765 truncated and 2839 full-length repnodes) and produced two large clusters (Figure 5.3A). The majority of singleton nodes are from metagenome sequencing or are candidate phyla. Cluster II (1181 repnodes) was composed almost solely of alpha and gamma proteobacterial species and primarily contained homogenous full-length sequences (Figure 5.3). Contrastingly, the largest cluster (Cluster I, 3914 repnodes) contained a mixture of both truncated and full-length sequences. The majority of full-length sequences in Cluster I belong to Bacteroidota and Bacillota phyla and the Betaproteobacterial Burkholderiales order Figure 5.4 and C.4). The shared clustering of these full-length isoforms with truncated isoforms indicates that these sequences share a greater amount of similarity, than the full-length predominantly proteobacterial isoforms in Cluster II, and might suggest that these sequences share a more recent common ancestor.

To examine the taxonomic distribution of truncated isoforms a second network was created using sequences that had a IPR042122 domain length between 62-77 residues (Figure 5.3B). The majority of the sequences cluster into a single group (Cluster A) with an edge similarity cutoff of 33%. Notably, all Thermotoga sequences formed an individual cluster, suggesting a greater degree of sequence dissimilarity from Cluster A, and is consistent with the deep-branching nature of this phylum. Additionally, Cluster A networks were created and refined with an edge cutoff of 68% (Figure 5.3C) and 72% (Figure 5.3D) and show the predominant phyla clusters are Bacillota (Cluster 1A), betaproteobacteria (Cluster 1B), Cyanobacteriota (Cluster 1C) and Actinomycetota (Cluster 1D). The majority of truncated isoforms sequences cluster by phyla, which is consistent with the vertical inheritance of *cysE*.

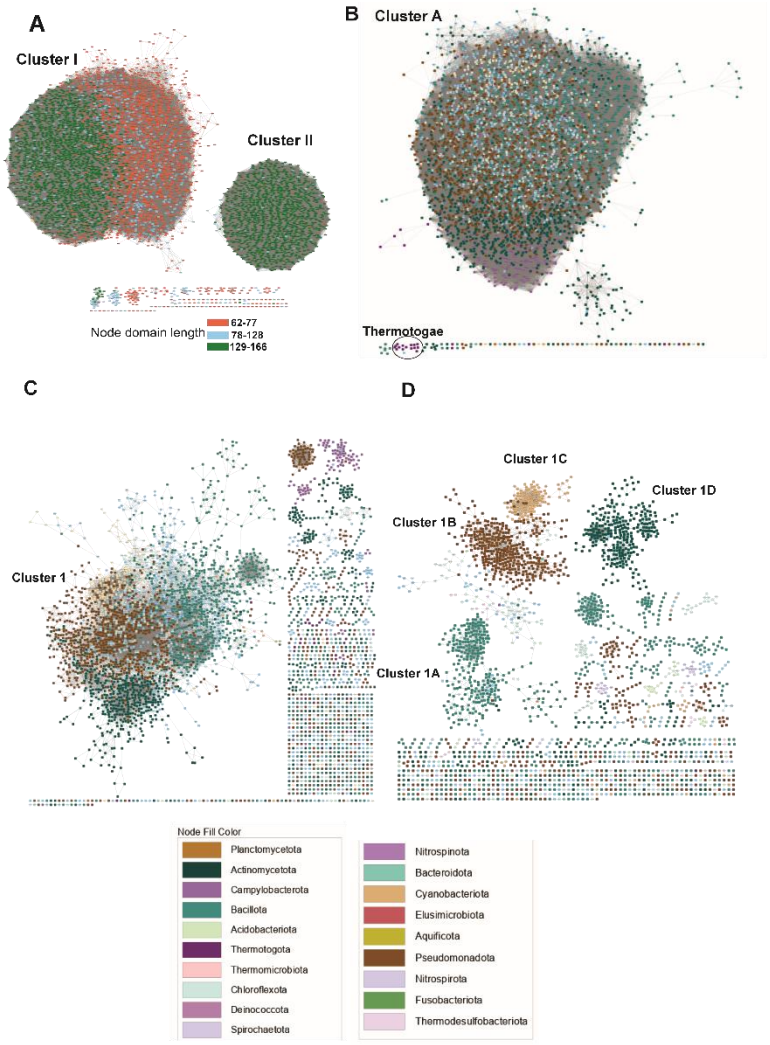


Figure 5.3 Repnode Sequence Similarity Network of bacterial CysE N-terminal SATase domain (IPR042122). (A) Sequence clusters have a sequence similarity edge cutoff of 17% (Repnode 60%). Sequences were filtered to have an SATase domain that was between 62-166 residues long to include both truncated and full-length isoforms. Nodes are coloured based on IPR042122 domain length, with truncated (62-77) and full-length (129-166) forms, coloured orange and green, respectively. Sequences that have an intermediate SATase domain length (78-128) are coloured blue. (B) Cluster A contains only truncated sequences (IPR042122 6277 residue length) and is refined to a 33% edge cutoff (Repnode 75%). (C) Network was generated from Cluster A and refined to a 68% edge cutoff (Repnode 75%). (D) Network was generated from Cluster 1 and refined to a 72% edge cutoff. Nodes are coloured by phylum for (B), (C) and (D) as show in legend. Figure was visualised using Cytoscape (Shannon et al., 2003).

5.4.3 CysE truncated isoforms are the prevalent isoform in Gram-positive species.

To visualise the phylogenetic relationship of CysE, sequences were mapped onto the iTOL tree (Ciccarelli et al., 2006). The iTOL tree was originally constructed from a concatenated alignment of 31 protein families for 191 species and represents a global phylogeny for the kingdoms of life (Ciccarelli et al., 2006). The bacterial genomes from the iTOL tree were searched with CysE protein sequences retrieved from UniProt (Figure 5.4). Sequences were determined to have a truncated or full length SATase domain based on the annotated length of the IPR042122 region (Table C.4).

The iTOl tree demonstrates widespread distribution of CysE isoforms across the bacterial kingdom. Of the 20 annotated phyla (based on iTOL annotation (Ciccarelli et al., 2006)) CysE is present in the following 16; Acidobacteriota, Actinomycetota, Bacteroidota, Chloroflexota, Chlorobiota, Cyanobacteriota, Fibrobacterota, Bacillota, Planctomycetota, all Proteobacteria (alpha, beta, gamma, delta and epsilon), Spirochaetota and Thermotogota. Absences were observed for iTOL groups Aquificota, Chlamydiota, Fusobacterota and Deinococcota. Given the lack of a *de novo* cysteine biosynthesis pathway for Chlamydiota members (Yang, Rajeeve, Rudel, & Dandekar, 2019), the absence of CysE is expected. Interestingly, both Deinococcota members, *Thermus thermophilus* and *Deinococcus radiodurans*, and the Aquificota species *Aquifex aeolicus*, did have an annotated CysK in their genomes (Ramoneda, Jensen, Price, Casamayor, & Fierer, 2023), but no CysE present. Additionally, several members of the *Lactobacillus* genus did not have an annotated CysE, but studies have shown that *metA* homoserine *O*-succinyl transferase from *Lactobacillus casei* (Bogicevic et al., 2016) and *Lactobacillus acidophilus* (Bastard et al., 2017) have serine *O*-acetyltransferase activity, and likely fulfils the role of CysE in these species. The full-length CysE isoform was the most prevalent isoform for Chloroflexota, Fibrobacterota, Planctomycetota, and Proteobacteria (alpha, beta, and gamma) except for delta and epsilon members (*Helicobacter spp.*, *Wolinella succinogenes* and *Campylobacter jejuni*). Correspondingly, the truncated isoform is the most prevalent isoform for the majority of iTOL bacterial phyla including Acidobacteriota, Actinomycetota, Cyanobacteriota, Bacillota, Spirochaetota, Thermotogota and epsilonproteobacteria.

A recent study by Gupta et al. (2022) examined CysE isoforms from the Mycobacteriaceae family and demonstrate presence of a truncated N-terminal domain, which is consistent with our findings reported here (Gupta & Gupta, 2022).

Multiple CysE isoforms are observed for plants (Yi et al., 2013) and protozoa (Kumar et al., 2011; Nozaki et al., 1999) with differences in sub-cellular localisation, kinetic efficiency and cysteine sensitivity seen between isoforms. All characterised bacterial isoforms to date exist as the sole copy of CysE in their genomes (Table C.1). The majority of iTOL species examined contained a single copy of *cysE*, irrespective of isoform type (Table C.4). Only a handful of species (8/151, 5.3%) had two or more *cysE* copies per genome including *Fibrobacter succinogenes*, *Thermotoga maritima*, *Gloeobacter violaceus*, *Nostoc* sp. PCC 7120, *Geobacter sulfurreducens*, *Rhodopseudomonas palustris*, *Bacteroides thetaiotaomicron* and *Synechococcus elongatus*. *Bacteroides thetaiotaomicron* and *S. elongatus* species contain both a full-length and truncated isoform. Intriguingly, the full-length CysE isoform harboured by *S. elongatus* is encoded on a plasmid (Figure 5.4). This isoform contains an extended SATase domain when compared to Proteobacterial isoforms (Table C.4). As seen for eukaryotes, multiple CysE copies within an organism have been shown to exhibit different biochemical characteristics (Yi et al., 2013) and could offer a fitness advantage for these bacteria, such as *S. elongatus*.

Overall, the distribution of truncated and full-length CysE isoforms across the iTOL tree species is consistent with SSN observations. One curiosity is the wide distribution of full-length isoforms for iTOL species from Planctomycetota, Fibrobacterota, Bacteroidota, Chlorobiota and proteobacteria phyla. To investigate the phylogenetic relationship between these full-length isoforms we constructed a maximum likelihood tree using a trimmed alignment of both the N-terminal SATase and the C-terminal LβH domain using a selection of CysE sequences from bacterial iTOL species (Figure 5.4) and PDB structures (Figure C.5). The Proteobacteria and the other four phyla form two distinct clades supported by a bootstrap value of 77 and 100, respectively. These two clades are consistent with the clustering seen in the SSN (Figure 5.3) and the phylogenetic relationship between these species seen in the iTOL tree (Figure 5.4) suggesting that these are the result of vertical inheritance and not recent horizontal gene transfer. Additionally, a

maximum-likelihood tree was constructed to investigate the relationship between truncated isoforms from annotated species present on iTOL tree (Bacillota, Actinomycetota, beta proteobacteria and epsilon proteobacteria), however we were unable to build a robust tree using truncated isoform sequences (Figure C.5), is likely due to sequence dissimilarity between these isoforms.

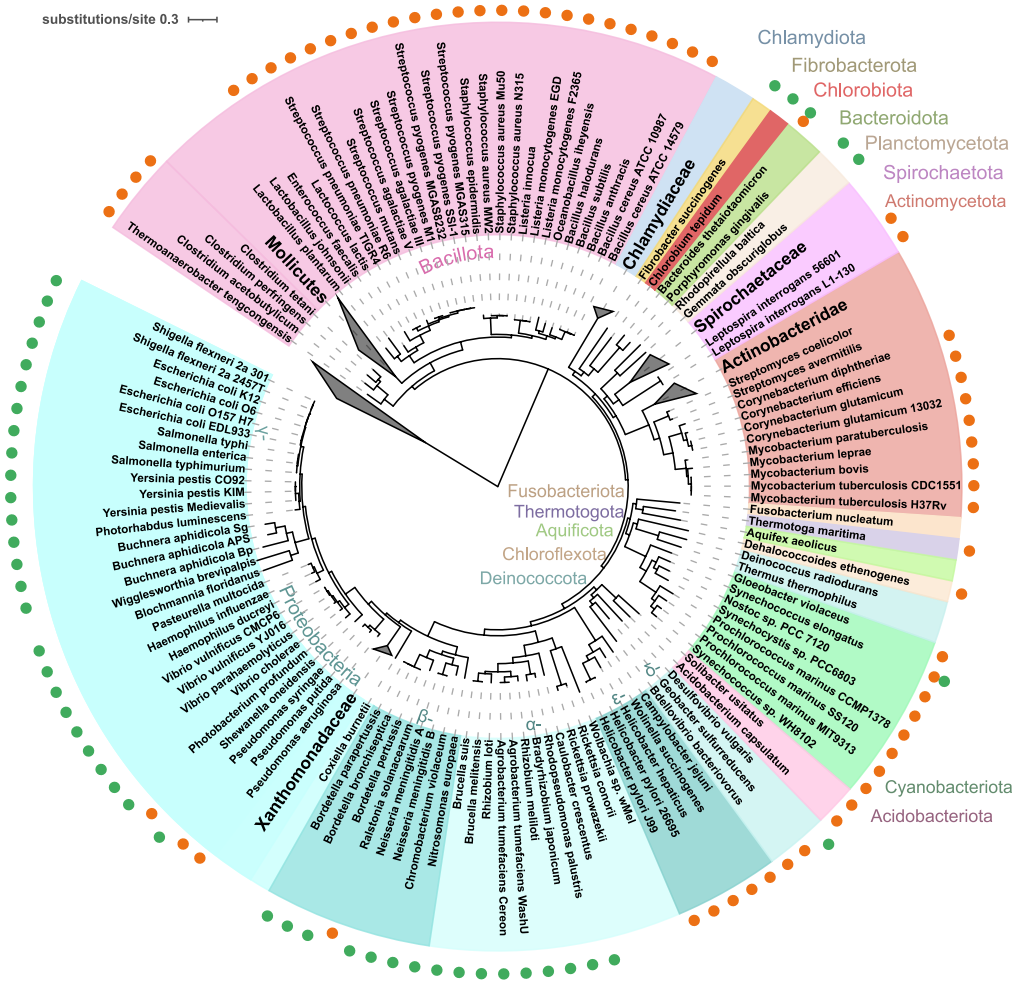


Figure 5.4 Annotation of iTOl tree bacterial species with CysE isoforms (Ciccarelli et al., 2006). The truncated and full-length isoforms are annotated with orange and circles, respectively. The 20 major bacterial phyla are annotated as in (Ciccarelli et al., 2006). CysE was encoded on the chromosome for all bacterial species except *Rhodopirellula baltica* and *Synechococcus elongatus* which harbours a plasmid encoded CysE. Bacterial species with no CysE isoform identified and the eukaryota/archaea domains are collapsed for clarity.

5.4.3 Truncated CysE isoforms are likely to be trimeric

To date, there is limited published biophysical characterisation of truncated CysE isoforms, raising the question; What oligomeric species do these isoforms adopt in solution? Recent advances in *in silico* modelling tools can provide insight into possible oligomeric species. Homology modelling of the truncated CysE monomer has been successfully used for screening inhibitors of *Sa*CysE (Chen et al., 2019) and *Mtb*CysE (Gupta & Gupta, 2021). Additionally, alignment of truncated and full-length monomer models shows strong structural conservation of N-terminal SATase and C-terminal LβH domains (Figure 5.1 and Figure 5.2). Here, we have used AlphaFold-Multimer (Evans et al., 2022; Jumper et al., 2021) via ColabFold (Mirdita et al., 2022) to determine the oligomeric states of *Sa*CysE.

Predictive modelling was used to build both trimeric and hexameric models for *Sa*CysE. For modelling the *Sa*CysE hexamer, six *Sa*CysE monomeric chains (1,278 residues) were supplied as input and three *Sa*CysE chains (639 residues) were supplied for modelling the *Sa*CysE trimer. The hexamer and trimer models with the highest confidence scores (predicted local distance difference test (pLDDT), predicted template modelling score (pTM) and interface predicted template modelling score (iPTM) were relaxed using AMBER (Hornak et al., 2006) and used for downstream analysis (Figure 5.5A and B). Ramachandran analysis of 1278 residues of the SaCysE hexamer showed 90.8% of residues in the favoured region, with 8.7% and 0.5% in the allowed and disallowed regions, respectively (Laskowski, MacArthur, Moss, & Thornton, 1993), indicating that the hexamer model has good stereochemistry. The hexamer additionally showed confident scoring (pLDDT 85.9, pTM 0.835, iPTM 0.829) for all chains (Figure C.6). The domains predicted with the highest confidence are N-terminal SATase and LβH domains (pLDDT > 90) (Figure 5.5A). The SaCysE hexamer is composed of two trimers, with three-fold symmetry and were orientated with the N-terminal SATase domains facing each other forming an apparent hexamer, however manual inspection of the hexameric interface revealed the presence of clashing non-bonded contacts between polar residues trimers, which remained despite repeated attempts at relaxation (Figure 5.5A). Conversely, the SaCysE trimer was predicted with greater confidence than the hexamer (pLDDT 87.4, PTM 0.839, iPTM 0.835) (Figure 5.5B) and inspection of the monomer-monomer interactions for the *Sa*CysE

trimer are similar to those seen in *Ec*CysE (0.632 RMSD for 403 α-carbons). Interestingly, the α-helical tails in both *Sa*CysE models interact with each other to form an α-helical coil, but this is predicted with low confidence and should be interpreted with caution.

The surface charge of a protein is critical for protein-protein interactions, with perturbations in charge being detrimental for complex formation. To investigate the clashing non-bonded contact SaCysE hexamer residues an electrostatic surface map was calculated for both the SaCysE hexamer and trimer. The N-terminal SATase domain surface of the SaCysE trimer involved in hexamer formation consists of a largely uncharged surface, which contrasts with the equivalent surface in EcCysE (Figure 5.5C). However, the *Sa*CysE trimer contains a series of positively charged residues that encircle the outward edge of the trimer surface, which are the same non bonded contact residues identified in the *Sa*CysE hexamer (Figure 5.5C). A positively charged trimer surface is also observed for the truncated *Mtb*CysE trimer (Gupta & Gupta, 2021). More recently, there has been experimental evidence to show that both trimer and hexamer configurations are seen in solution for the truncated isoform *Mtb*CysE (Rahisuddin et al., 2024). In the absence of a crystal/cryo-EM structure, the authors used AlphaFold and were only able to model the *Mt*CysE trimer, consistent with our *Sa*CysE modelling. Therefore, we predict that *Sa*CysE adopts a trimer in solution, but formation of hexamer is a possibility and ultimately needs to be confirmed through experimental characterisation.

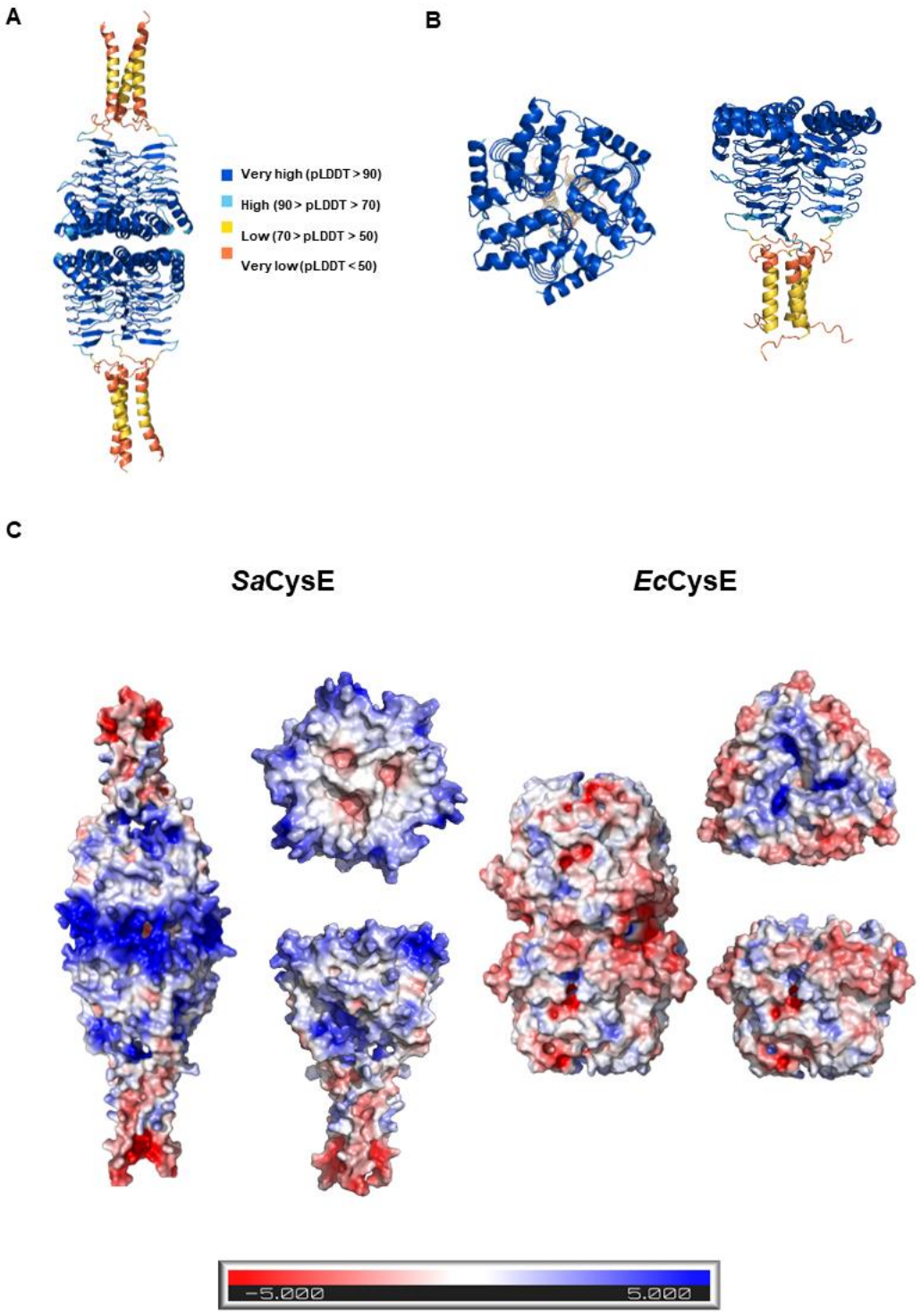


Figure 5.5 Electrostatic map of predicted multimeric model of *Sa*CysE (P6776). (A) ColabFold predicted *Sa*CysE hexamer. (B) ColabFold predicted *Sa*CysE trimer. Model is coloured by residue confidence scoring predicted local distance difference test (pLDDT); high confidence (blue, pLDDT > 90), confident (light blue, 90 > pLDDT > 70), low confidence (yellow, 70 > pLDDT > 50) and very low confidence

(orange, pLDDT < 50). (C) The electrostatic for *Sa*CysE and *Ec*CysE hexamers and trimers (PDB 1T3D). Protein surface is coloured red to blue (-5 Kt/e to 5 Kt/e).

In summary, using predictive modelling we demonstrate that *Sa*CysE is likely to form a trimer in solution. Given the importance of the CysE hexamer for cysteine synthase complex formation (Benoni et al., 2017), it is unlikely that the *Sa*CysE trimer can form the CSC with CysK from *S. aureus*. Conversely, recent characterisation of the truncated *Mtb*CysE showed that CSC formation was observed for the *Mtb*CysE hexamer, but unexpectedly the complex led to a reduction in activity of both *Mtb*CysE and *Mtb*CysK, which is at odds with all published CSC data to date ((Benoni et al., 2017; Kumaran, Yi, Krishnan, & Jez, 2009; Mino, Yamanoue, et al., 2000)). This observation may reflect the presence of alternative pathways for cysteine biosynthesis that can bypass CysE in *M. tuberculosis* (Rahisuddin et al., 2024). These findings suggest that *Sa*CysE might be able to form the cysteine synthase complex, given presence of the critical C-terminal residue is an isoleucine (Figure 5.1) (Mino, Hiraoka, et al., 2000), but is reliant whether it can form a hexamer. Ultimately, experimental characterisation of *Sa*CysE is needed to observe the oligomeric states adopted in solution and its ability to form the cysteine synthase complex, which will in turn help elucidate the role of these truncated isoforms in cysteine biosynthesis.

5.5 Conclusions

Serine acetyltransferases catalyse the first step in the *de novo* biosynthesis of cysteine and act as a key point of pathway regulation. Here, we report that bacterial CysE isoforms can be categorised into two distinct groups, defined by the presence of a full-length or a truncated N-terminal SATase domain. Furthermore, we report that this truncation is structurally discrete, with the loss shortening the N-terminal domain from eight to four α-helices. We demonstrate that the truncated isoform is the predominant isoform amongst Gram-positive bacteria, while full-length isoforms dominate Proteobacteria. Through structural analysis we predicted the truncation would favour formation of a trimer. Overall, based on these findings it is clear that the loss of the N-terminal SATase impacts the function and regulation of the truncated isoform, but it is unclear what the evolutionary advantage of this isoform is and additional experimental characterisation is needed to elucidate the

biochemical differences between full-length and truncated isoforms and their respective roles in *de novo* cysteine biosynthesis.

5.6 Materials and Methods

5.6.1 Multiple Sequence Alignments

Sequences were retrieved from UniProt (https://www.uniprot.org/) (Consortium, 2022) for all bacterial CysE isoforms deposited to the RSCB Protein Data Bank (PDB) (https://www.rcsb.org/). The MAFFT alignment tool was used for generating MSA (Katoh & Standley, 2013) and MSA figures were created using ESPript (3.0) (Robert & Gouet, 2014).

5.6.2 Sequence Similarity Networks

Sequence similarity networks (SSNs) were created using the Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST) server (database version 2023_03/97, UniProt 2023-05) (Oberg, Zallot, & Gerlt, 2023; Zallot, Oberg, & Gerlt, 2019). The initial network was generated using the families option (option B) using the InterPro domain IPR042122 (N-terminal SATase domain) and UniRef90 cluster IDs (39,897 "Full size" vs. 15,804 UniRef90 sequences) with alignment performed over the length of the IPR042122 domain. UniProt-defined fragments were excluded. The network was taxonomically filtered to include only bacterial sequences. The SSN edge E-value calculation was set to 5, with a protein family size fraction of 1. For the Repnode 60 network, the alignment score was set to 17. Finalisation of the SSN for examining domain length were filtered to include sequences with a IPR042122 domain between 62 and 166 residues in length. For the Repnode 75 network for analysing sequence taxonomy the initial alignment score of 23 was used. Due to computational limitations, a representative node network was used for all analysis. Visualisation of all networks was performed using Cytoscape (version 3.9.1) (Shannon et al., 2003) and was used to set the sequence similarity edge cut-off.

5.6.3 Generation of iTOL tree and Phylogenetic Tree Building for CysE sequences for iTOL species

For annotation of CysE isoforms to the iTOL tree (Ciccarelli et al., 2006), the CysE peptide sequences were retrieved from UniProt for iTOL bacterial species

(Ciccarelli et al., 2006) and confirmed to have an annotated N-terminal SATase domain (IPR042122). Length of the SATase domain was recorded and confirmed through manually inspection of predicted models from the AlphaFold database (https://alphafold.ebi.ac.uk/, accessed January 2024). For species with no CysE sequence present in UniProt, a combination of BLASTP (The National Center for Biotechnology Information, NCBI) and manual inspection of the bacterial genomes using NCBI genome viewer (accessed January 2024) was used to confirm the absence. The iTOl figure was generated using the iTOl server (Letunic & Bork, 2021).

The phylogenetic tree was created for all CysE sequences that were found in iTOL tree bacterial species. BLASTP with CysE from *N. gonorrhoeae* as a query sequence was used to search against bacterial iTOL tree species in the NCBI database. Where present, a CysE sequence was retrieved for each iTOL species. The protein sequences containing the N-terminal SATase domain (IPR042122) for building the CysE phylogenetic tree were retrieved from UniProt (Consortium, 2022) and multiple whole sequence alignment was performed using MAFFT (multiple alignment using fast Fourier transform) (Katoh & Standley, 2013). Alignments were visualised using Geneious (2020.0.5) and trimmed manually removing the C-terminal alpha helical tails (downstream of LβH domain), with the trimmed sequences containing both the SATase domain (IPR042122) and the conserved LβH domain (IPR045304) for tree building (Capella-Gutiérrez, SillaMartínez, & Gabaldón, 2009). IQ-TREE (Nguyen, Schmidt, von Haeseler, & Minh, 2014) was used for tree building and model selection using ModelFinder (Kalyaanamoorthy, Minh, Wong, von Haeseler, & Jermiin, 2017), and ultrafast bootstrap approximation (UFBoot) (Minh, Nguyen, & von Haeseler, 2013) were used for calculating bootstrap values. The best model IQ-TREE was accessed using the Galaxy Server (Galaxy Version 2.1.2, https://usegalaxy.org.au/) (Community, 2022). Tree figures were generated using iTOL tree website (version 5) (Letunic & Bork, 2021).

5.6.4 ColabFold and AlphaFold2 Models

Predicted protein models were built using a local ColabFold (version 1.5.2) installation (https://github.com/YoshitakaMo/localcolabfold) on a NVIDIA

GeForce RTX 3060 (Mirdita et al., 2022). Sequence alignments were generated using mmseqs2_uniref_env in unpaired_paired mode. Templates were used in prediction (PDB100) and models were built with five recycles. For multimeric prediction the model alphafold2_multimer_v3 was used. For building the *Sa*CysE hexamer six CysE amino acid chains were supplied as input and three chains were supplied for building the *Sa*CysE trimer. All models were relaxed using Assisted Model Building with Energy Refinement (AMBER) as a part of ColabFold (Case et al., 2005). PROCHECK (Laskowski et al., 1993) was used to perform Ramachandran analysis of models and was accessed via the PDBsum server (Laskowski, Jabłońska, Pravda, Vařeková, & Thornton, 2018). Electrostatics were calculated for ColabFold models using the Adaptive Poisson-Boltzmann Solver (APBS) plugin in PyMOL using default settings (Jurrus et al., 2018).

5.7 Supplementary Material

Supplementary Materials associated with Chapter Five can be found in Appendix C.

5.8 Future Directions

This work provides the first in-depth analysis of serine acetyltransferases (SAT/CysE) across bacteria and provides insight into the distribution and diversity of CysE isoforms, whereas previous research and characterisation has focused on full-length CysE isoforms (Benoni et al., 2017; Hindson & Shaw, 2003; Pye et al., 2004). This work supports further bioinformatic and experimental investigation of truncated CysE isoforms to elucidate functional and regulatory differences and how this impacts their role in *de novo* cysteine biosynthesis.

Given the limited characterisation of the truncated isoform, one of the key questions is; what are the consequences of this truncation on the quaternary structure of CysE. Previous work has performed kinetic characterisation of the truncated isoforms, however limited biophysical data is published at time of writing with only the characterisation of CysE from *M. tuberculosis* (as discussed above) (Rahisuddin et al., 2024). Our research group is expanding into the characterisation of the *de novo* cysteine biosynthesis pathway in *Staphylococcus aureus*, as this organism has lost the sulphate reduction arm but has retained a CysK and truncated

CysE (*Sa*CysE) homologue (Lithgow, Hayhurst, Cohen, Aharonowitz, & Foster, 2004), which draw parallels to the cysteine biosynthesis pathway present in *N. gonorrhoeae* (Hicks & Mullholland, 2018). As an extension of this thesis, I have conducted preliminary work into the biophysical characterisation of *Sa*CysE. The planned experiments are to investigate the oligomeric state and ability to form the cysteine synthase complex (CSC) using analytical size exclusion chromatography, and use X-ray crystallography to determine the *Sa*CysE crystal structure. Additionally, although kinetic data is available for truncated isoforms (Chen et al., 2019; Qiu et al., 2014; Qiu et al., 2013; Rahisuddin et al., 2024) regulation by cysteine has not been determined. Given the presence of an extended C-terminal α-helical tail seen for truncated isoforms (Figure 5.1 and Figure 5.2), and the key role of the tail in feedback inhibition as seen in cysteine bound crystal structures (Pye et al., 2004; Toyomoto et al., 2023), this extended tail could impact the sensitivity to cysteine inhibition. To measure cysteine inhibition, dose-response and mode of inhibition curves will be collected using a direct 232 nm method that is already established for measuring CysE activity (Oldham et al., 2022). Together, this work would provide the first biophysical characterisation of *Sa*CysE and the first characterisation of cysteine inhibition of a truncated CysE isoform.

Preliminary work has already been completed for *Sa*CysE characterisation. Our cloning and purification methods were adapted from Chen and colleagues (Chen et al., 2019). Cloning of CysE from *S. aureus* MW2 (BA000033.2:561091-561732, 639 nucleotides, 213 residues, ordered as a synthetic gene) into pET28b-PstI using restriction sites PstI and XhoI to express the *Sa*CysE with a C-terminal Histag. Cloning was verified using sequencing and transformed into *E. coli* BL21 (DE3) for protein expression. However, despite repeated optimisation attempts, including recloning with a N-terminal Histag, temperature and detergent screens, efforts to express and purify soluble *Sa*CysE have been unsuccessful. Given that previous work by Chen et al. (2019) demonstrated soluble *Sa*CysE expression is possible, cloning into different plasmids and transforming into *E. coli* expression strains better suited for insoluble proteins will form part of future purification efforts.

The wide distribution of the truncated CysE isoform is an unexpected finding, and raises the question; what is the evolutionary history of serine acetyltransferases? As mentioned previously, CysE is a member of the left-handed β-helix family of

acyltransferases (Raetz & Roderick, 1995), however to date there is no literature examining the evolution of CysE enzymes. In this thesis, we have highlighted the existence of both a full-length and truncated CysE isoform, however it is unclear what the ancestral form resembled and the evolutionary trajectory that resulted in two structurally distinct isoforms. Based on phylogenetic work in this chapter, the use of Bayesian interference methods could be used to reconstruct the ancestral state of CysE. The success of this strategy may be constrained by low sequence similarity between truncated and full-length isoforms. Alternatively, determining the origin of the full-length isoform is a better candidate for investigation, given the higher degree of sequence conservation between full-length species. Additionally, research has revealed an unexpectedly high degree of sequence and structural similarity between plant and full-length bacterial isoforms (Yeon, Yoo, Takagi, & Kang, 2018). It is possible that this similarity is the result of a horizontal gene transfer event, but further bioinformatic investigation, using plant and a broader selection of bacterial CysE sequences is needed to determine to determine this.

5.9 References

Bastard, K., Perret, A., Mariage, A., Bessonnet, T., Pinet-Turpault, A., Petit, J.-L., . . . de Berardinis, V. (2017). Parallel evolution of non-homologous isofunctional enzymes in methionine biosynthesis. *Nature Chemical Biology, 13*(8), 858-866. 10.1038/nchembio.2397

Benoni, R., De Bei, O., Paredi, G., Hayes, C. S., Franko, N., Mozzarelli, A., . . . Campanini, B. (2017). Modulation of Escherichia coli serine acetyltransferase catalytic activity in the cysteine synthase complex. *FEBS letters, 591*(9), 1212-1224. 10.1002/18733468.12630

Bogicevic, B., Berthoud, H., Portmann, R., Bavan, T., Meile, L., & Irmler, S. (2016). Cysteine biosynthesis in Lactobacillus casei: identification and characterization of a serine acetyltransferase. *FEMS Microbiol Lett, 363*(4) 10.1093/femsle/fnw012

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics, 25*(15), 1972-1973. 10.1093/bioinformatics/btp348

Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., . . . Woods, R. J. (2005). The Amber biomolecular simulation programs. *J Comput Chem, 26*(16), 1668-1688. 10.1002/jcc.20290

Chen, C., Yan, Q., Tao, M., Shi, H., Han, X., Jia, L., . . . Ma, Y. (2019). Characterization of serine acetyltransferase (CysE) from methicillin-resistant Staphylococcus aureus and inhibitory effect of two natural products on CysE. *Microbial Pathogenesis, 131*, 218-226. 10.1016/j.micpath.2019.04.002

Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., & Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science, 311*(5765), 1283-1287. 10.1126/science.1123061

Community, T. G. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research, 50*(W1), W345-W351. 10.1093/nar/gkac247

Consortium, T. U. (2022). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research, 51*(D1), D523-D531. 10.1093/nar/gkac1052 Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., . . . Hassabis, D. (2022). Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034. 10.1101/2021.10.04.463034

Gagnon, Y., Breton, R., Putzer, H., Pelchat, M., Grunberg-Manago, M., & Lapointe, J. (1994). Clustering and co-transcription of the Bacillus subtilis genes encoding the aminoacyl-tRNA synthetases specific for glutamate and for cysteine and the first enzyme for cysteine biosynthesis. *J Biol Chem, 269*(10), 7473-7482.

Gorman, J., & Shapiro, L. (2004). Structure of serine acetyltransferase from Haemophilus influenzae Rd. *Acta Crystallogr D Biol Crystallogr, 60*(Pt 9), 1600-1605. 10.1107/s0907444904015240

Gupta, S., & Gupta, V. (2021). Homology modeling, structural insights and in-silico screening for selective inhibitors of mycobacterial CysE. *J Biomol Struct Dyn, 39*(5), 1547-1560. 10.1080/07391102.2020.1734089

Gupta, S., & Gupta, V. (2022). Unveiling the structural features of CysE: a novel target for therapeutic interventions against persistent mycobacteria. *Journal of Applied and Natural Science, 14*(2), 531 - 542. 10.31018/jans.v14i2.3461

Hicks, J. L., & Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology, 164*(12), 1471-1480. 10.1099/mic.0.000728

Hicks, J. L., Oldham, K. E. A., McGarvie, J., & Walker, E. J. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. *Biosci Rep, 42*(10) 10.1042/bsr20220368

Hindson, V. J. (2003). Serine acetyltransferase of Escherichia coli: substrate specificity and feedback control by cysteine. *The Biochemical journal, 375*(Pt 3), 745-752. 10.1042/BJ20030429

Hindson, V. J., & Shaw, W. V. (2003). Random-Order Ternary Complex Reaction Mechanism of Serine Acetyltransferase from Escherichia coli. *Biochemistry, 42*(10), 3113-3119. 10.1021/bi0267893

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution, 35*(2), 518-522. 10.1093/molbev/msx281

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins, 65*(3), 712-725. 10.1002/prot.21123

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589. 10.1038/s41586-021-03819-2

Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., . . . Baker, N. A. (2018). Improvements to the APBS biomolecular solvation software suite. *Protein Science, 27*(1), 112-128. 10.1002/pro.3280

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods, 14*(6), 587-589. 10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution, 30*(4), 772-780. 10.1093/molbev/mst010

Kim, Y., Zhou, M., Peterson, S., Anderson, W. F., & Joachimiak, A. (2006). *Crystal Structure of Serine Acetyltransferase CysE from Yersinia pestis*.

Kitabatake, M., So, M. W., Tumbula, D. L., & Söll, D. (2000). Cysteine biosynthesis pathway in the archaeon Methanosarcina barkeri encoded by acquired bacterial genes? *J Bacteriol, 182*(1), 143-145. 10.1128/jb.182.1.143-145.2000

Kredich, N. (2008). Biosynthesis of Cysteine. *EcoSal Plus* doi:10.1128/ecosalplus.3.6.1.11 Kredich, N. M., & Tomkins, G. M. (1966). The Enzymic Synthesis of l-Cysteine in Escherichia coli and Salmonella typhimurium. *Journal of Biological Chemistry, 241*(21), 4955-4965.

Kumar, N., Singh, R. P., & Kumaran, S. (2019). *Understanding Mechanics of competitive allostery Using Engineered Cysteine Synthase Assembly*.

Kumar, S., Kumar, N., Alam, N., & Gourinath, S. (2014). Crystal structure of serine acetyl transferase from Brucella abortus and its complex with coenzyme A. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1844*(10), 1741-1748. 10.1016/j.bbapap.2014.07.009

Kumar, S., Raj, I., Nagpal, I., Subbarao, N., & Gourinath, S. (2011). Structural and Biochemical Studies of Serine Acetyltransferase Reveal Why the Parasite Entamoeba histolytica Cannot Form a Cysteine Synthase Complex. *Journal of Biological Chemistry, 286*(14), 12533-12541. 10.1074/jbc.M110.197376

Kumaran, S., Yi, H., Krishnan, H. B., & Jez, J. M. (2009). Assembly of the cysteine synthase complex and the regulatory role of protein-protein interactions. *J Biol Chem, 284*(15), 10268-10275. 10.1074/jbc.M900154200

Kuzin, A. P., Chen, Y., Seetharaman, J., Sahdev, S., Wang, D., Mao, L., . . . Tong, L. (2008). *Three dimensional structure of the serine acetyltransferase from Bacteroides vulgatus, NORTHEAST STRUCTURAL GENOMICS CONSORTIUM TARGET BVR62*.

Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein Sci, 27*(1), 129-134. 10.1002/pro.3289

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography, 26*(2), 283-291. 10.1107/S0021889892009944

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research, 49*(W1), W293W296. 10.1093/nar/gkab301

Lithgow, J. K., Hayhurst, E. J., Cohen, G., Aharonowitz, Y., & Foster, S. J. (2004). Role of a cysteine synthase in Staphylococcus aureus. *Journal of bacteriology, 186*(6), 1579-1590. 10.1128/jb.186.6.1579-1590.2004

Minh, B. Q., Nguyen, M. A. T., & von Haeseler, A. (2013). Ultrafast Approximation for Phylogenetic Bootstrap. *Molecular Biology and Evolution, 30*(5), 1188-1195. 10.1093/molbev/mst024

Mino, K., Hiraoka, K., Imamura, K., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2000). Characteristics of serine acetyltransferase from Escherichia coli deleting different lengths of amino acid residues from the C-terminus. *Biosci Biotechnol Biochem, 64*(9), 1874-1880. 10.1271/bbb.64.1874

Mino, K., Yamanoue, T., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2000). Effects of bienzyme complex formation of cysteine synthetase from Escherichia coli on some properties and kinetics. *Biosci Biotechnol Biochem, 64*(8), 1628-1640. 10.1271/bbb.64.1628

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods, 19*(6), 679-682. 10.1038/s41592-022-01488-1

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2014). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood

Phylogenies. *Molecular Biology and Evolution, 32*(1), 268-274. 10.1093/molbev/msu300

Nozaki, T., Asai, T., Sanchez, L. B., Kobayashi, S., Nakazawa, M., & Takeuchi, T. (1999). Characterization of the Gene Encoding Serine Acetyltransferase, a Regulated Enzyme of Cysteine Biosynthesis from the Protist Parasites Entamoeba histolytica and Entamoeba dispar: Regulation and possible function of the cysteine biosynthetic pathway in Enatmoeba. *Journal of Biological Chemistry, 274*(45), 32445-32452. 10.1074/jbc.274.45.32445

Oberg, N., Zallot, R., & Gerlt, J. A. (2023). EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools. *Journal of Molecular Biology, 435*(14), 168018. 10.1016/j.jmb.2023.168018

Oldham, K. E. A., Prentice, E. J., Summers, E. L., & Hicks, J. L. (2022). Serine acetyltransferase from Neisseria gonorrhoeae; structural and biochemical basis of inhibition. *Biochem J, 479*(1), 57-74. 10.1042/bcj20210564

Olsen, L., B., H., Vetting, M., & Roderick, S. (2004). Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor. *Biochemistry, 43*(20), 6013-6019.

Park, S., & Imlay, J. A. (2003). High Levels of Intracellular Cysteine Promote Oxidative DNA Damage by Driving the Fenton Reaction. *Journal of Bacteriology, 185*(6), 1942-1950. 10.1128/jb.185.6.1942-1950.2003

Pye, V. E., Tingey, A. P., Robson, R. L., & Moody, P. C. E. (2004). The Structure and Mechanism of Serine Acetyltransferase from Escherichia coli. *Journal of Biological Chemistry, 279*(39), 40729-40736. 10.1074/jbc.M403751200

Qiu, J., Ma, Y., Owusu, L., Jiang, T., & Xin, Y. (2014). Functional analysis of serine acetyltransferase from Mycobacterium smegmatis. *J Basic Microbiol, 54*(7), 670677. 10.1002/jobm.201300858

Qiu, J., Wang, D., Ma, Y., Jiang, T., & Xin, Y. (2013). Identification and characterization of serine acetyltransferase encoded by the Mycobacterium tuberculosis Rv2335 gene. *International Journal of Molecular Medicine, 31*, 1229-1233. 10.3892/ijmm.2013.1298

Raetz, C. R. H., & Roderick, S. L. (1995). A Left-Handed Parallel β Helix in the Structure of UDP-N-Acetylglucosamine Acyltransferase. *Science, 270*(5238), 997-1000. 10.1126/science.270.5238.997

Rahisuddin, R., Thakur, P., Kumar, N., Saini, N., Banerjee, S., Singh, R. P., . . . Kumaran, S. (2024). Multi-oligomeric and catalytically compromised serine acetyltransferase and cysteine regulatory complex of Mycobacterium tuberculosis. *Biochimie, 221*, 110-124. 10.1016/j.biochi.2024.01.009

Ramoneda, J., Jensen, T. B. N., Price, M. N., Casamayor, E. O., & Fierer, N. (2023). Taxonomic and environmental distribution of bacterial amino acid auxotrophies. *Nature Communications, 14*(1), 7608. 10.1038/s41467-023-43435-4

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research, 42*(W1), W320-W324. 10.1093/nar/gku316

Sauerwald, A., Zhu, W., Major, T. A., Roy, H., Palioura, S., Jahn, D., . . . Söll, D. (2005). RNA-dependent cysteine biosynthesis in archaea. *Science, 307*(5717), 1969-1972. 10.1126/science.1108329

Seattle Structural Genomics Center for Infectious Disease (SSGCID), Abendroth, J., Gardberg, A., & Staker, B. (2010). *Crystal structure of WW/RSP5/WWP domain: bacterial transferase hexapeptide repeat: serine O-Acetyltransferase from Brucella Melitensis*.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: a software environment for integrated models of

biomolecular interaction networks. *Genome Res, 13*(11), 2498-2504. 10.1101/gr.1239303

Tarique, K. F., Abdul Rehman, S. A., & Gourinath, S. (2012). *Crystal structure of Serine acetyltransferase from Vibrio cholerae O1 biovar El Tor N16961*.

Toyomoto, T., Ono, K., Shiba, T., Momitani, K., Zhang, T., Tsutsuki, H., . . . Sawa, T. (2023). Alkyl gallates inhibit serine O-acetyltransferase in bacteria and enhance susceptibility of drug-resistant Gram-negative bacteria to antibiotics. *Frontiers in Microbiology, 14* 10.3389/fmicb.2023.1276447

Vaara, M. (1992). 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*(3), 249-254. 10.1111/j.1574-6968.1992.tb05471.x

Verma, D., Gupta, S., Saxena, R., Kaur, P., R, R., Srivastava, S., & Gupta, V. (2020). Allosteric inhibition and kinetic characterization of Klebsiella pneumoniae CysE: An emerging drug target. *Int J Biol Macromol, 151*, 1240-1249. 10.1016/j.ijbiomac.2019.10.170

Yang, M., Rajeeve, K., Rudel, T., & Dandekar, T. (2019). Comprehensive Flux Modeling of Chlamydia trachomatis Proteome and qRT-PCR Data Indicate Biphasic Metabolic Differences Between Elementary Bodies and Reticulate Bodies During Infection. *Frontiers in Microbiology, 10* 10.3389/fmicb.2019.02350

Yeon, J. Y., Yoo, S. J., Takagi, H., & Kang, H. A. (2018). A Novel Mitochondrial Serine O-Acetyltransferase, OpSAT1, Plays a Critical Role in Sulphur Metabolism in the Thermotolerant Methylotrophic Yeast Ogataea parapolymorpha. *Scientific Reports, 8*(1), 2377. 10.1038/s41598-018-20630-8

Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B., & Jez, J. M. (2013). Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. *The Journal of biological chemistry, 288*(51), 3646336472. 10.1074/jbc.M113.527143

Zallot, R., Oberg, N., & Gerlt, J. A. (2019). The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry, 58*(41), 41694182. 10.1021/acs.biochem.9b00735

Chapter Six: Conclusion and Future Perspectives

6.1 Thesis Summary

Neisseria gonorrhoeae is the etiological agent of the sexually transmitted infection, gonorrhoea. The rapid emergence of antimicrobial-resistant strains has resulted in resistance to every class of antibiotic used for frontline treatment (Unemo & Shafer, 2011). The emergence of ceftriaxone-resistant strains has resulted in it being labelled a priority pathogen by the World Health Organisation, highlighting the urgent demand for new antimicrobials (WHO, 2017). The absence of the *de novo* cysteine biosynthesis pathway in humans provides an excellent opportunity for antimicrobial exploration (Campanini et al., 2015). Progress into identifying inhibitors for CysE and CysK have been made for both Gram-negative and Gram-positive pathogens. This thesis demonstrates my work to characterise and use structure-based virtual inhibitor screening for the cysteine biosynthetic enzyme, serine acetyltransferase from *Neisseria gonorrhoeae* (*Ng*CysE).

We hypothesised that *Ng*CysE was a promising target for antimicrobial development, given the stressful host environment (Seib et al., 2006) and the essentiality of *cysE* (Remmele et al., 2014). As *Ng*CysE was uncharacterised, the collection of kinetic and structural data for *Ng*CysE was essential for using structure-based virtual inhibitor screening to identify *Ng*CysE inhibitors. To allow *Ng*CysE characterisation, methods for expression and purification of *Ng*CysE were optimised. Using our purification methods, we were able to produce *Ng*CysE of sufficient purity and yield for downstream assays and analysis.

Given the disruption of the sulphate acquisition pathway for cysteine biosynthesis in *N. gonorrhoeae* (Hicks & Mullholland, 2018), kinetic characterisation was needed to ascertain that *Ng*CysE had serine acetyltransferase activity. To characterise the kinetics of *Ng*CysE, Michaelis-Menten activity assays and inhibition assays were collected. The Michaelis-Menten assays demonstrated *Ng*CysE had comparable activity to characterised bacterial homologues. Through the collection of inhibitory dose response and mode of inhibition assays we show

that *Ng*CysE is sensitive to feedback inhibition by L-cysteine, and show that cysteine is a competitive inhibitor relative to both substrates; serine and acetyl-CoA. Overall, these results show that CysE from *N. gonorrhoeae* is a functional serine acetyltransferase that is inhibited by cysteine and is comparable with other bacterial homologues, demonstrating that this enzyme is an active serine acetyltransferase and likely contributes to the synthesis of cysteine in *N. gonorrhoeae.*

The oligomeric state of serine acetyltransferases is a known determinant of activity, with hexamer exhibiting increased activity over trimers (Rahisuddin et al., 2024). Using analytical size exclusion chromatography (SEC), we show that *Ng*CysE elutes as a hexamer. These findings are supported by SEC-small angle X-ray scattering (SAXS) data where only a *Ng*CysE hexamer species is observed with no smaller oligomeric species present. Overall, we show that *Ng*CysE exists solely as hexamer in solution which is consistent with characterised full-length CysE homologues.

To identify *Ng*CysE inhibitors we employed the strategy of structure-based virtual inhibitor screening, which required a high-quality atomic structure of *Ng*CysE. Using recombinant *Ng*CysE an optimal crystallisation condition was identified resulting in the collection of a 2.01Å dataset. Model refinement revealed *Ng*CysE had crystallised as a homohexamer with 3:2 fold symmetry. The hexamer was assembled through the interaction of two trimers via the N-terminal α-helical SATase domain. Each trimer was formed through interactions of both the N-terminal α-helical SATase domain and C-terminal left-handed β helix domain. *Ng*CysE active sites are located in clefts between neighbouring monomers in the trimer, resulting in six active sites per hexamer. Intriguingly, we saw extensive density for the C-terminal tails, which are often absent in crystal structures given their flexibility (Gorman & Shapiro, 2004; Olsen, B., Vetting, & Roderick, 2004; Toyomoto et al., 2023; Yi et al., 2013), as a result of domain swapping between adjacent hexamers in the crystal lattice. Examination of *Ng*CysE revealed supporting density for active site residues and the presence of L-malate, which was present in the crystallisation condition. Crystallisation of a second *Ng*CysE structure with L-serine bound, reveals shared interactions of both L-malate and L-serine, but only L-serine interacts with catalytic dyad residue His162, which

reflects that L-serine is a native substrate of *Ng*CysE. Overall, we have shown that *Ng*CysE has a hexameric oligomeric structure consistent with serine acetyltransferases and our L-malate structure is of sufficient quality for downstream inhibitor screening.

To conduct structure-based virtual inhibitor screening the *Ng*CysE structure had to be prepared for library screening. The *Ng*CysE structure underwent model relaxation and water removal. The screening of a commercially available drug-like and a large molecular compound library was conducted, producing a total of 48 hit compounds of which 28 were ordered for screening. To test for inhibitory activity *in vitro,* a DTNB-coupled assay was adapted for high throughput plate screening of inhibitors against *Ng*CysE. The *in vitro* screening of inhibitors demonstrated the best inhibitor was compound 2 producing an IC_{50} of 13.9 µM, making it one of the most potent CysE inhibitors published to date. Analysis of docking interactions between compound 2 and *Ng*CysE shows key hydrogen bonding interactions along the length of the compound with both acetyl-CoA and L-serine binding pocket residues. Overall, this research presents the first characterisation and inhibitors of *Ng*CysE, and is also the first published attempt to target a sulphur metabolism pathway in any *Neisseria* species. This thesis demonstrates that the targeting of the *de novo* cysteine biosynthesis pathways is a promising source of new antimicrobial targets *N. gonorrhoeae.*

In addition, we have extensively reviewed research into the antimicrobial targeting of serine acetyltransferases from bacterial pathogens. Upon performing this review, the presence of a unique N-terminal truncation for CysE homologues from *Mycobacterium tuberculosis* and *Staphylococcus aureus* (*Sa*CysE) was noted. The distribution and diversity of CysE isoforms for bacteria is unknown and as such we set out to answer this using bioinformatic analysis and predictive modelling tools. In our paper we set out to answer the following questions: (1) does the truncated isoform appear in other bacteria and is this truncation consistent, (2) what is the phylogenetic relationship between truncated and full-length CysE isoforms and (3) what are the consequences of this truncation on the oligomeric state of CysE. We show that truncated CysEs are widespread and are the predominant form for Gram-positive bacteria. Through generating sequence similarity networks for CysE we show that truncated isoforms bear more similarity to full-length isoforms

from Gram-positive bacteria than Proteobacteria. Predictive modelling suggests that the truncation would prevent likely assembling of a *Sa*CysE canonical hexamer, however we could predict confidently that *Sa*CysE can assemble as a trimer. Overall, this work is fundamental in establishing the diversity and distribution of the truncated CysE isoforms in bacteria, and the physiological consequences of this truncation for CysE enzymes.

6.2 Future Perspectives

Future perspective sections are available at the end of each chapter under the heading "Future Directions". Additional Future Directions are detailed in the following sections.

6.2.1 In vivo characterisation of cysE in Neisseria gonorrhoeae

As mentioned in the previous section we have biochemically characterised *Ng*CysE and show it to be a functional serine acetyltransferase *in vitro*, however *in vivo* characterisation is crucial aspect of determining the role of *Ng*CysE in *N. gonorrhoeae*. As a part of future work, we plan to examine the effects of CysE depletion on *N. gonorrhoeae* using a *cysE* knockdown strain. The essential nature of *cysE* (Remmele et al., 2014) has meant that it took an extended amount of time to create a viable knockdown strain (of which still needs to be optimised) and was only achieved at the end of this thesis (as discussed in Future Directions of Chapter Four Section 4.7.2). However, we would like to discuss here the future experiments that will be carried out using a *cysE* knockdown strain to determine the fundamental role of *Ng*CysE in both cysteine biosynthesis and pathogenicity.

The cysteine biosynthesis pathway has been shown to be upregulated during cell invasion for *Neisseria meningitidis* (Takahashi, Watanabe, Kim, Yokoyama, & Yanagisawa, 2018), however it is unknown if this extends to *N. gonorrhoeae*. To measure pathogenicity, the *cysE* knockdown strain would be subjected to both epithelial gonococcal cell invasion assays and biofilm formation assays. Additionally, cysteine biosynthesis is key for mitigating oxidative stress and conferring antibiotic resistance (Turnbull & Surette, 2010). To examine the role of *cysE*, the knockdown strain will be subjected to hydrogen peroxide assays and antibiotic sensitivity MIC assays. Moreover, testing if the antibiotic sensitivity

increases for a *cysE* depleted strain is additional validation that *Ng*CysE inhibitors could also act as antimicrobial adjuvant. Overall, this characterisation is crucial for strengthening the suitability of CysE for inhibition and given the key regulation of the *de novo* cysteine biosynthesis pathway occurs through the CysE catalytic step, the *in vivo* characterisation of the *cysE* knockdown strain will provide insight into the role of *de novo* cysteine biosynthesis pathway and elucidate the role of cysteine in *N. gonorrhoeae*.

The essentiality of *cysE* is uncommon for bacteria with disruption often resulting in cysteine auxotrophy (Fiegler & Brückner, 1997; Fujishima et al., 2018; Hitchcock, Kelly, Hitchcock, & Taylor, 2023), which raises questions over the essentiality of *cysE* in *N. gonorrhoeae*. As mentioned previously, we hypothesise that *N. gonorrhoeae* is reliant on *de novo* cysteine biosynthesis to meet high cysteine requirements, that maintains high concentrations of glutathione. Alternatively, CysE could have a moonlighting function in *N. gonorrhoeae* which makes *cysE* essential. Functions outside of cysteine biosynthesis have not been documented for CysE enzymes, however moonlighting functions in contact-dependent growth inhibition (CDI) have been reported for OASS (Johnson et al., 2016). To answer these questions growth, transcriptomic and metabolomic analysis of a CysE depleted *N. gonorrhoeae* using a *cysE* knockdown strain would reveal insights into pathways that are disrupted by CysE depletion and would highlight if CysE has an alternative function in *N. gonorrhoeae*. To do this, these experiments would be conducted using chemically defined media with either cysteine/cystine, or sulphide, supplied as the sole sulphur source. Any unique transcriptional or metabolomic differences seen for the *cysE* depleted strain compared to wildtype on cysteine rich media as a sulphur source would indicate ulterior functions in alternative cellular pathways.

6.2.2 Further characterisation and use of medicinal chemistry to develop potent *Ng*CysE inhibitors

As discussed in Chapter Four, performing minimal inhibitor testing (MIC) testing of compound 2 against *N. gonorrhoeae* is crucial for drug optimisation of this compound. Given that *cysE* is likely essential or conditionally essential based on our findings, if compound 2 is permeable to the bacterial membrane, then it is

likely we will see an MIC. Ideally the MIC should be in the same range as the inhibitor (13.9 µM), with a lower MIC suggesting off-target inhibition. Additionally, MIC testing should be conducted on chemically defined media with cysteine/cystine or sulphide as the sole sulphur source, as previous research demonstrated that media richness was associated with increased MICs (Magalhães et al., 2020). MIC testing of compound 2 will be collected on both rich Gonococcal Broth (GCB) media and chemically defined Graver-Wade media with different sulphur sources. This will allow us to examine the influence of different sulphur sources on MICs.

Off-target inhibition is a major hurdle in drug discovery and is seen for CysE antimicrobials (Hicks, Oldham, McGarvie, & Walker, 2022). Extensive work by Toyomoto and colleagues (2023) (as mentioned previously in Chapter 2) identified alkyl gallates as inhibitors of CysE, and while compound testing demonstrated the compounds were not cytotoxic, comparison of intracellular amino acid concentrations of gallate treated *E. coli* with a *cysE* knockout strain, revealed there were discrepancies in amino acid concentrations suggesting off-target inhibition. In our work (Chapter Five), to avoid off-target inhibition we excluded compounds during virtual screening that would likely be false positives using a pan-assay interference compounds (PAINS) filter. This filter excluded compounds with known PAINS structures (chemical structures know for interfering with assay (i.e. reacting with the components or covalently modifying the enzyme). Ultimately, future work characterising CysE inhibitors will also include cytotoxicity studies and metabolomic analysis, as an extra measure to screen out inhibitors with off-target inhibition.

6.2.3 6.2.3 The targeting of the *de novo* cysteine biosynthesis pathway in the pathogen *Neisseria meningitidis*

Neisseria meningitidis is the causative agent of invasive meningococcal disease. This often fatal disease results in bacterial meningitis or septicaemia with global fatality rate of 12% (Wang, Santoreneos, Giles, Haji Ali Afzali, & Marshall, 2019). Although meningococcal vaccines are available, increasing rates of antimicrobial resistance have been reported for invasive *N. meningitidis* isolates (Willerton et al., 2021). The *de novo* cysteine biosynthesis pathway has been implicated in

pathogenesis for *N. meningitidis*, having been shown to be upregulated during cell invasion (Takahashi et al., 2018) and in invasive *N. meningitidis* isolates (Ren et al., 2017). As highlighted by Hicks et al. (2018), despite having a high degree of genomic similarity (90%) for *N. gonorrhoeae* and *N. meningitidis*, a key difference is the absence of sulphate reduction pathway for *N. gonorrhoeae* and its presence in *N. meningitidis*. This is supported by growth experiments that demonstrated only *N. gonorrhoeae* was unable to grow on sulphate as a sole sulphur source (Le Faou, 1984). Given the essentiality of *cysE* in *N. meningitidis* (Capel et al., 2016; Muir et al., 2020), CysE from *N. meningitidis* (*Nm*CysE) is a potential target for antimicrobial development. Given that the cysteine biosynthetic enzymes of *N. meningitidis* are uncharacterised, the study of *Nm*CysE is a natural continuation of this thesis.

CysE kinetic assays already established as part of this thesis, would be used to determine the Michaelis-Menten kinetics and cysteine sensitivity of *Nm*CysE. Additionally, X-ray crystallography will be used to determine the atomic structure of *Nm*CysE and SEC-SAXS will be used to determine the oligomeric state. Given the high degree of similarity between *Ng*CysE and *Nm*CysE, *Ng*CysE inhibitors identified in this thesis could be tested against *Nm*CysE. Furthermore, given the essential nature of *cysE* for *N. meningitidis* (Capel et al., 2016; Muir et al., 2020) to study the role of *cysE in vivo* would be through the creation of a *cysE* knockdown strain. Given the differences in the sulphur assimilation pathways between these species, this would allow examination of the effect of an intact vs disrupted sulphate reduction pathway has on CysE function.

6.2.4 Characterisation of the cysteine synthase complex

The cysteine synthase complex (CSC) is key regulatory aspect of the *de novo* cysteine biosynthesis pathway in allowing cysteine synthesis to be tuned to sulphur availability (Benoni et al., 2017). Past attempts to form the CSC from *N. gonorrhoeae* (*Ng*CSC) using *Ng*CysE and CysK from *N. gonorrhoeae* (*Ng*CysK) enzymes have been unsuccessful (Oldham, 2020). Currently, PhD student Jack McGarvie (McGarvie, 2021) is working on optimisation of the *Ng*CSC complex. Given that *Ng*CysE has all characteristics needed for CSC formation, including purifying as a hexamer and the presence of a C-terminal isoleucine, we assume the

two enzymes can form a complex. However, a positive control is needed. Based on review of the literature, the CSC from *Escherichia coli* (*Ec*CSC) was selected as our positive control. Efforts made collectively by PhD student Jack McGarvie, MSc student Jessicca Usu and myself, have resulted in the successful purification of recombinant *E. coli* enzymes CysE (*Ec*CysE) and CysK (*Ec*CysK) and have successfully formed the *Ec*CSC. The formation was confirmed via size exclusion chromatography, SDS-PAGE and (SEC-SAXS) (unpublished results that are part of PhD thesis for Jack McGarvie), replicating results seen in the literature (Benoni et al., 2017). Additionally, the cysteine synthase complex has been shown to readily form through the combining of purified enzymes (Benoni et al., 2017) and is supported by our observation that recombinant *Ec*CysE or *Ec*CysK enzymes copurify with expression host-derived, *Ec*CysK or *Ec*CysE, respectively. The establishment of *Ec*CSC as a positive control in our lab is crucial for the optimisation of *Ng*CSC formation. As mentioned previously, presence of the C-terminal isoleucine and a hexameric assembly would permit *Ng*CysE to form the CSC, which is now supported by our ability to form the *Ec*CSC. Based on these findings, *Ng*CSC optimisation will be conducted as a part of Jack McGarvie's thesis where he will screen the addition of sulphur compounds that could be needed for complex formation.

6.3 Concluding Statement

The results presented in thesis have enhanced our understanding of serine acetyltransferases and their role in *de novo* cysteine biosynthesis. Our bioinformatic analysis demonstrated that the truncated SATase domain identified initially in *S. aureus* and *M. tuberculosis* is in fact widely distributed amongst bacterial phyla. This will hopefully lead to future in depth characterisation of this enigmatic truncated isoform. Our findings also deepen our understanding of *de novo* cysteine biosynthesis in *N. gonorrhoeae*. The biochemical characterisation of *Ng*CysE was fundamental for not only understanding its role in cysteine biosynthesis, but also instrumental in our discovery of the first *Ng*CysE inhibitors. This thesis is a testament to the antimicrobial potential of CysE inhibition and paves the way for CysE targeting in other bacterial pathogens.

6.4 References

Benoni, R., De Bei, O., Paredi, G., Hayes, C. S., Franko, N., Mozzarelli, A., . . . Campanini, B. (2017). Modulation of Escherichia coli serine acetyltransferase catalytic activity in the cysteine synthase complex. *FEBS letters, 591*(9), 1212-1224. 10.1002/18733468.12630

Campanini, B., Pieroni, M., Raboni, S., Bettati, S., Benoni, R., Pecchini, C., . . . Mozzarelli, A. (2015). Inhibitors of the Sulphur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Current Medicinal Chemistry, 22*(2), 187-213. 10.2174/0929867321666141112122553

Capel, E., Zomer, A. L., Nussbaumer, T., Bole, C., Izac, B., Frapy, E., . . . Coureuil, M. (2016). Comprehensive Identification of Meningococcal Genes and Small Noncoding RNAs Required for Host Cell Colonization. *mBio, 7*(4), e01173-01116. 10.1128/mBio.01173-16

Fiegler, H., & Brückner, R. (1997). Identification of the serine acetyltransferase gene of Staphylococcus xylosus. *FEMS Microbiol Lett, 148*(2), 181-187. 10.1111/j.15746968.1997.tb10286.x

Fujishima, K., Wang, K. M., Palmer, J. A., Abe, N., Nakahigashi, K., Endy, D., & Rothschild, L. J. (2018). Reconstruction of cysteine biosynthesis using engineered cysteine-free enzymes. *Scientific Reports, 8*(1), 1776. 10.1038/s41598-018-19920y

Gorman, J., & Shapiro, L. (2004). Structure of serine acetyltransferase from Haemophilus influenzae Rd. *Acta Crystallogr D Biol Crystallogr, 60*(Pt 9), 1600-1605. 10.1107/s0907444904015240

Hicks, J. L., & Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology, 164*(12), 1471-1480. 10.1099/mic.0.000728

Hicks, J. L., Oldham, K. E. A., McGarvie, J., & Walker, E. J. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. *Biosci Rep, 42*(10) 10.1042/bsr20220368

Hitchcock, N., Kelly, D. J., Hitchcock, A., & Taylor, A. J. (2023). Cysteine Biosynthesis in Campylobacter jejuni: Substrate Specificity of CysM and the Dualism of Sulphide. *Biomolecules, 13*(1), 86.

Johnson, P. M., Beck, C. M., Morse, R. P., Garza-Sánchez, F., Low, D. A., Hayes, C. S., & Goulding, C. W. (2016). Unraveling the essential role of CysK in CDI toxin activation. *Proceedings of the National Academy of Sciences, 113*(35), 9792-9797. doi:10.1073/pnas.1607112113

Le Faou, A. (1984). Sulphur nutrition and metabolism in various species of Neisseria. *Annales de l'Institut Pasteur / Microbiologie, 135*(1, Supplement B), 3-11. 10.1016/S0769-2609(84)80037-X

Magalhães, J., Franko, N., Raboni, S., Annunziato, G., Tammela, P., Bruno, A., . . . Costantino, G. (2020). Inhibition of Nonessential Bacterial Targets: Discovery of a Novel Serine O-Acetyltransferase Inhibitor. *ACS Medicinal Chemistry Letters* 10.1021/acsmedchemlett.9b00627

McGarvie, J. (2021). *Structural and biochemical characterisation of O-acetylserine sulphydrylase (CysK) from Neisseria gonorrhoeae* (Masters). The University of Waikato, Hamilton, New Zealand. Retrieved from https://hdl.handle.net/10289/14596

Muir, A., Gurung, I., Cehovin, A., Bazin, A., Vallenet, D., & Pelicic, V. (2020). Construction of a complete set of Neisseria meningitidis mutants and its use for the phenotypic profiling of this human pathogen. *Nature Communications, 11*(1), 5541. 10.1038/s41467-020-19347-y

Oldham, K. (2020). *Cysteine biosynthesis and the role of CysE in Neisseria gonorrhoeae* (Masters). The University of Waikato, Hamilton, New Zealand. Retrieved from https://hdl.handle.net/10289/13461

Olsen, L., B., H., Vetting, M., & Roderick, S. (2004). Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor. *Biochemistry, 43*(20), 6013-6019.

Rahisuddin, R., Thakur, P., Kumar, N., Saini, N., Banerjee, S., Singh, R. P., . . . Kumaran, S. (2024). Multi-oligomeric and catalytically compromised serine acetyltransferase and cysteine regulatory complex of Mycobacterium tuberculosis. *Biochimie, 221*, 110-124. 10.1016/j.biochi.2024.01.009

Remmele, C. W., Xian, Y., Albrecht, M., Faulstich, M., Fraunholz, M., Heinrichs, E., . . . Rudel, T. (2014). Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic acids research, 42*(16), 10579-10595. 10.1093/nar/gku762

Ren, X., Eccles, D. A., Greig, G. A., Clapham, J., Wheeler, N. E., Lindgreen, S., . . . MacKichan, J. K. (2017). Genomic, Transcriptomic, and Phenotypic Analyses of Neisseria meningitidis Isolates from Disease Patients and Their Household Contacts. *mSystems, 2*(6), e00127-00117. 10.1128/mSystems.00127-17

Seib, K. L., Wu, H.-J., Kidd, S. P., Apicella, M. A., Jennings, M. P., & McEwan, A. G. (2006). Defenses against oxidative stress in Neisseria gonorrhoeae: a system tailored for a challenging environment. *Microbiology and molecular biology reviews : MMBR, 70*(2), 344-361. 10.1128/MMBR.00044-05

Takahashi, H., Watanabe, H., Kim, K. S., Yokoyama, S., & Yanagisawa, T. (2018). The Meningococcal Cysteine Transport System Plays a Crucial Role in Neisseria meningitidis Survival in Human Brain Microvascular Endothelial Cells. *mBio, 9*(6), e02332-02318. 10.1128/mBio.02332-18

Toyomoto, T., Ono, K., Shiba, T., Momitani, K., Zhang, T., Tsutsuki, H., . . . Sawa, T. (2023). Alkyl gallates inhibit serine O-acetyltransferase in bacteria and enhance susceptibility of drug-resistant Gram-negative bacteria to antibiotics. *Frontiers in Microbiology, 14* 10.3389/fmicb.2023.1276447

Turnbull, A. L., & Surette, M. G. (2010). Cysteine biosynthesis, oxidative stress and antibiotic resistance in Salmonella typhimurium. *Research in Microbiology, 161*(8), 643-650. 10.1016/j.resmic.2010.06.004

Unemo, M., & Shafer, W. M. (2011). Antibiotic resistance in Neisseria gonorrhoeae: origin, evolution, and lessons learned for the future. *Ann N Y Acad Sci, 1230*, E19-28. 10.1111/j.1749-6632.2011.06215.x

Wang, B., Santoreneos, R., Giles, L., Haji Ali Afzali, H., & Marshall, H. (2019). Case fatality rates of invasive meningococcal disease by serogroup and age: A systematic review and meta-analysis. *Vaccine, 37*(21), 2768-2782. 10.1016/j.vaccine.2019.04.020

WHO. (2017). World Health Organization (WHO) Global Priority List of AntibioticResistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics. Genenva: WHO.

Willerton, L., Lucidarme, J., Walker, A., Lekshmi, A., Clark, S. A., Walsh, L., . . . Borrow, R. (2021). Antibiotic resistance among invasive Neisseria meningitidis isolates in England, Wales and Northern Ireland (2010/11 to 2018/19). *PLoS One, 16*(11), e0260677. 10.1371/journal.pone.0260677

Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B., & Jez, J. M. (2013). Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. *The Journal of biological chemistry, 288*(51), 3646336472. 10.1074/jbc.M113.527143

Appendix A: Supplementary Material for Chapter Three

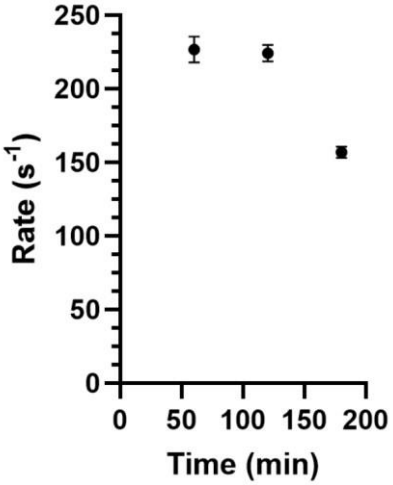


Figure A.1. *Ng*CysE activity versus time post-purification. *Ng*CysE was stored at room temperature post-purification and activity measured every 60 minutes (x-axis). Assays were collected with 0.625 µg *Ng*CysE, 1 mM serine and 0.45 mM acetyl-CoA. Data points are mean and error bars are SEM derived from two replicates.

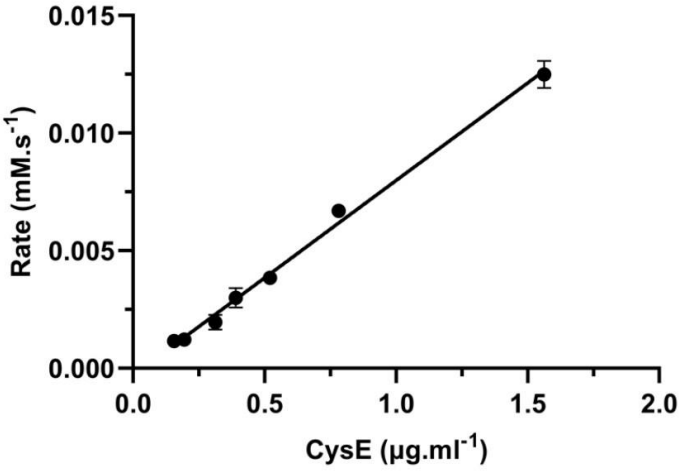


Figure A.2 *Ng*CysE activity measured at varying *Ng*CysE concentrations. Activity was measured for 0.156, 0.312, 0.391, 0.521, 0.781 and 1.56 µg.ml^{-1} of NgCysE (x-axis). Initial rates were measured in assay buffer in the presence of 0.45 mM acetyl-CoA and 10 mM L-serine. Enzyme stocks were stored at room temperature (22°C) for the duration of the assays. Data points are means derived from duplicates and standard error is displayed as error bars.

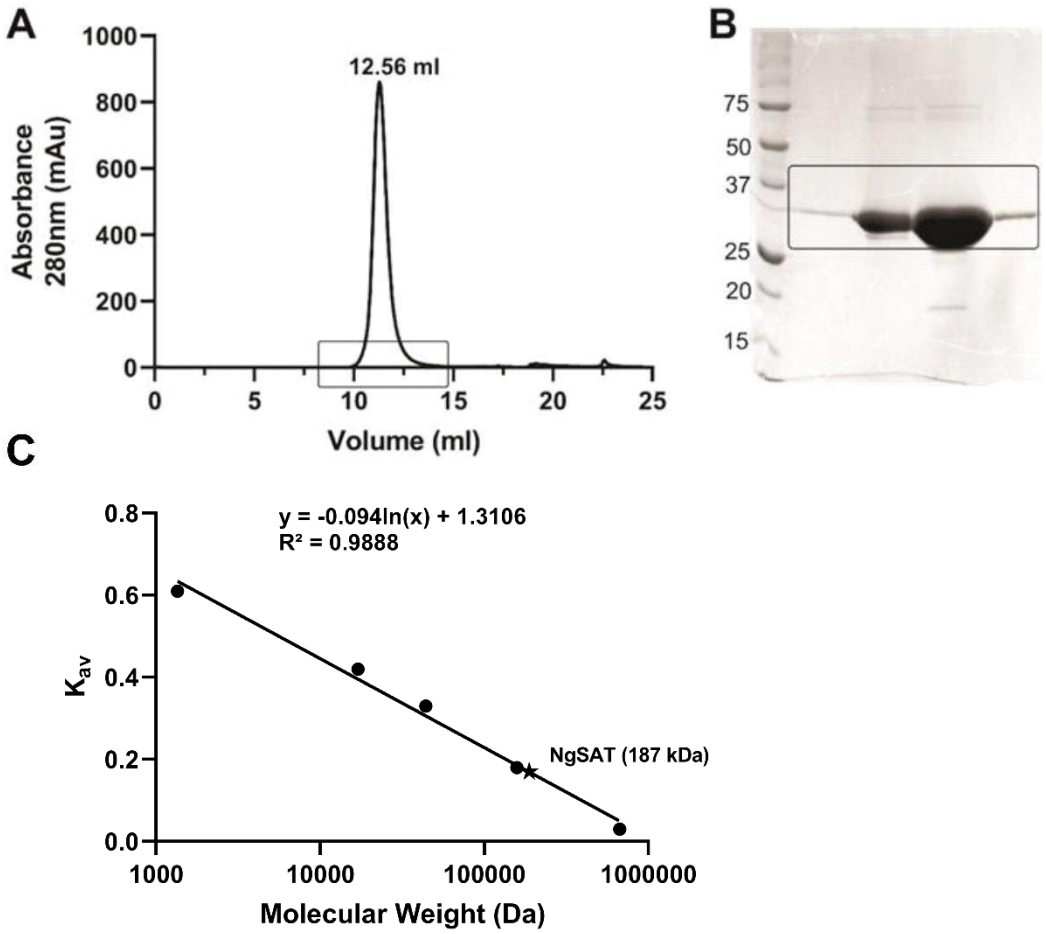


Figure A.3 Gel filtration purification of *Ng*CysE. (A) Gel filtration chromatogram of purified *Ng*CysE using an Enrich 650 Gel Filtration Column (Bio-Rad) with a single elution peak at 12.56 ml. (B) Corresponding 12% SDS-PAGE gel of column fractions (outlined by a box on the chromatogram and the SDS-PAGE gel) showing a high yield and purity of *Ng*CysE. Molecular weights of Precision Plus Protein Standards (Bio-rad) in kDa are labelled. (C) Calibration of the Enrich 650 Gel Filtration Column. Molecular weights of gel filtration calibration standards (Bio-Rad) plotted on the x-axis. Kav plotted on the y-axis. Kav of *Ng*CysE (0.17) indicated by star. Kav calculated by subtracting column void volume from the protein elution volume and dividing by the total column volume. Equation from line of best fit used for determining molecular weight of unknown proteins.

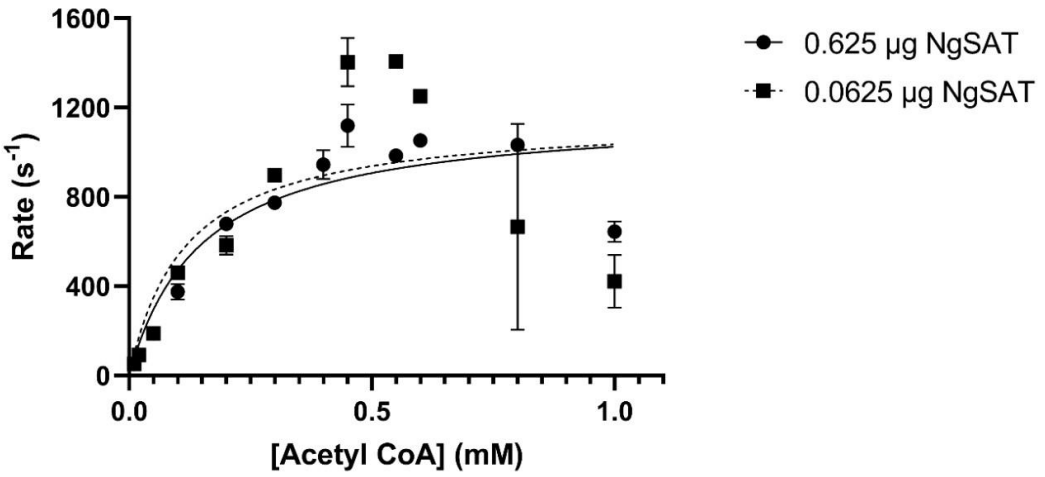


Figure A.4 Kinetic analysis of *Ng*CysE substrate acetyl-CoA with two different concentrations of *Ng*CysE (0.625 and 0.0625 µg) at 10 mM L-serine. Although reproducible at two different concentrations, with two different enzyme preparations, data for the 10-fold lower *Ng*CysE concentration (0.0625 µg) show increased error due to fluctuations in the data at higher starting absorbance values due to high acetyl-CoA concentrations. The Michaelis-Menten equation is fit (solid line represents 0.625 µg *Ng*CysE (R^2 = 0.8449) and dashed line represents 0.0625 µg *Ng*CysE (R^2 = 0.5240)) for both data sets and shows reproducibility. The overall fit for the equation is reasonable, with the exception of data collected for 1 mM acetyl-CoA which displays inhibition that is likely substrate inhibition. Due to limitations in experimental setup, we were unable to collect *Ng*CysE rates for higher acetyl-CoA concentrations and as such do not feel comfortable fitting a substrate inhibition model due to only one point (1 mM acetyl-CoA) displaying inhibition. Enzyme concentration was factored into rates by dividing by the enzyme concentration to give the rate (s^{-1}). Plotted data points represent mean alongside SEM of two replicates.

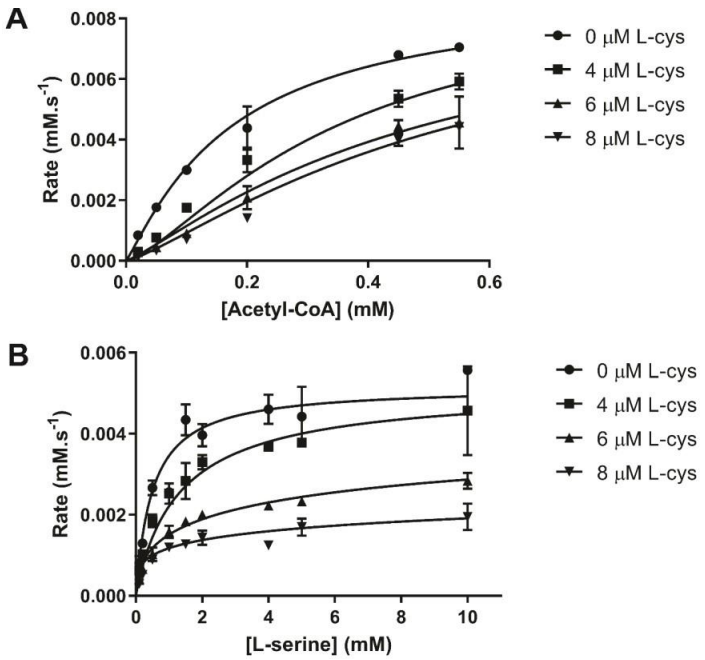


Figure A.5 Characterisation of cysteine inhibition, incorporation of the Hill coefficient to account for cooperativity. Data were collected for 0 (circles), 4 (squares), 6 (triangles) and 8 (inverted triangles) µM of L-cysteine. (A) Rates versus acetyl-CoA concentration plotted, with line representing competitive inhibition equation with incorporation of the hill coefficient. (B) Rates versus L-serine concentration plotted, with line representing competitive inhibition equation with incorporation of the hill coefficient. Data points are mean and error bars are SEM derived from two replicates.

Equation A.1 Supplementary Equation. [S] = substrate inhibition, [I] = inhibitor concentration and h = hill co-efficient.

$$Rate = \frac{V_{max}[S]^h}{[S]^h + K_M{}^h(1 + \frac{[I]}{K_i})}$$

Table A.1 Competitive Inhibition + Hill coefficient values.

	0 µM L-cys	**4 µM L-cys**	**6 µM L-cys**	**8 µM L-cys**
Acetyl-CoA				
h[a]	1.137 ± 0.1507	1.405 ± 0.1162	1.203 ± 0.1127	1.275 ± 0.1080
Ki[b] (µM)	2.473 ±0.5823	2.473 ±0.5823	2.473 ±0.5823	2.473 ±0.5823
R-squared	0.9851	0.9746	0.9791	0.9804
L-serine				
h[a]	1.005 ± 0.140	0.949 ± 0.081	0.481 ± 0.0553	0.298 ± 0.059
Ki[b] (µM)	2.508 ± 0.330	2.508 ± 0.330	2.508 ± 0.330	2.508 ± 0.330
R-squared	0.9030	0.9096	0.9610	0.8035

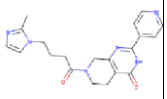
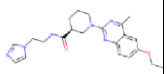
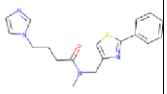
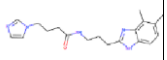
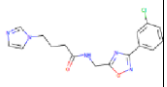
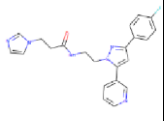
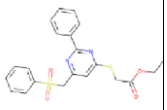
[a]Hill coefficient < 1 indicates negatively cooperative binding ~ 1 indicates non-cooperative binding >1 indicates positively cooperative binding.

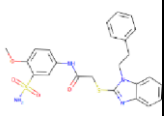
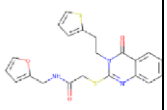
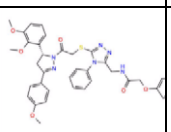
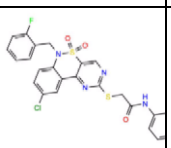
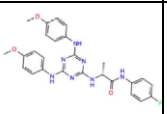
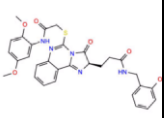
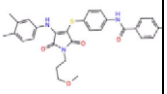
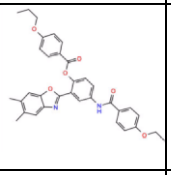
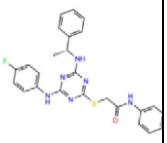
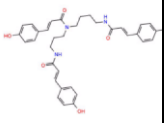
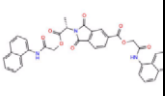
[b]Inhibitory constant (Ki)

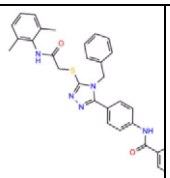
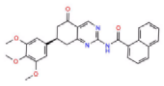
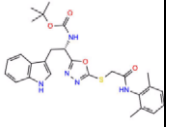
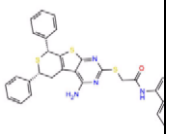
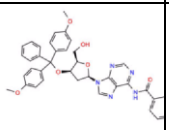
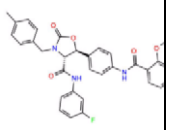
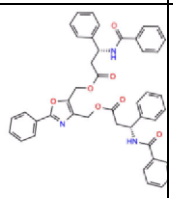
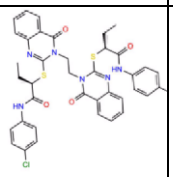
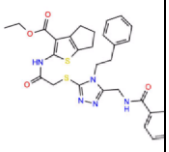
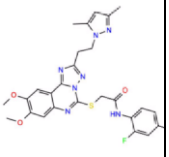
Error represents SEM of two replicates.

Appendix B: Supplementary Material for Chapter Four

Table B.1 List of inhibitors identified by virtual screening and tested experimentally. Lower MM-GBSA energy indicates greater ligand affinity. *Compounds that produced a negative rate had percent remaining activity rounded to 0%.

No.	Compound	Structure	Docking score (kcal/mol)	MM-GBSA binding energy (kcal/mol)	Remaining % activity at 100 µM (Vi/Vo)	Solubility in 100 µM screen
1	ZINC72123855		-9.342	-73.89	40.7	Soluble
2	ZINC11882369		-9.708	-60.5	20.6	Soluble
3	ZINC72459177		-9.867	-65.55	21.2	Soluble
4	ZINC244690946		-9.3	-60.41	26.8	Soluble
5	ZINC71908845		-9.647	-60.12	4.7	Soluble
6	ZINC19340516		-9.548	-63.11	62.0	Soluble
7	ZINC1383061		-8.561	-57.63	6.7	Insoluble

8	ZINC9492101		-8.870	-51.46	17.8	Insoluble
9	ZINC57473376		-7.7763	-57.55	45.8	Insoluble
10	ZINC150410062		-8.4	-82.2	52.26	Insoluble
11	ZINC102452962		-8.4	-87.4	172	Insoluble
12	ZINC9908138		-8.4	-78.1	99.7	Insoluble
13	ZINC102745388		-8.5	-93.4	152.8	Soluble
14	ZINC103389177		-8.7	-83.8	0* (-6.49)	Insoluble
15	ZINC8384491		-8.7	-80.4	0* (-12.0)	Soluble
16	ZINC680526		-8.7	-62.2	24.7	Insoluble
17	ZINC56874950		-8.8	-91.4	233.3	Soluble
18	ZINC150392050		-8.8	-60.8	0*(-93.5)	Soluble

19	ZINC8438183		-8.8	-86.7	23.9	Insoluble
20	ZINC2212876		-8.9	-73.3	93.3	Soluble
21	ZINC2220593		-9.0	-68.0	78.8	Soluble
22	ZINC1844763		-9.0	-62.5	9.4	Insoluble
23	ZINC95722620		-9.2	-79.9	6.5	Insoluble
24	ZINC103168280		-9.3	-80.4	141.83	Soluble
25	ZINC97995347		-9.4	-84.7	86.19	Insoluble
26	ZINC150443539		-9.5	-77.1	29.6	Insoluble
27	ZINC101654223		-9.6	-96.5	13.97	Insoluble
28	ZINC33356340		-10.8	-81.4	157.7	Soluble

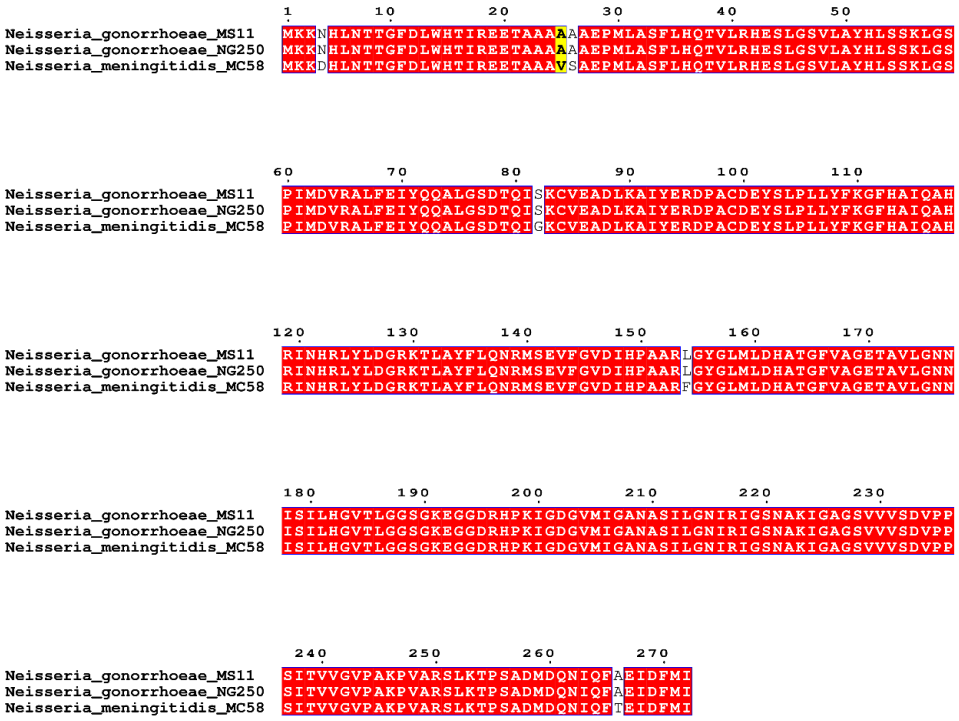


Figure B.1 Multiple sequence alignment of the CysE sequences from pathogenic *Neisseria*. *N. gonorrhoeae* NG250 CysE has a single nucleotide mutation (silent) and *N. gonorrhoeae* MS11 is representative of 156/157 CysE *N. gonorrhoeae* sequences in NCBI. CysE from *N. meningitidis* MC58 has been included for comparison.

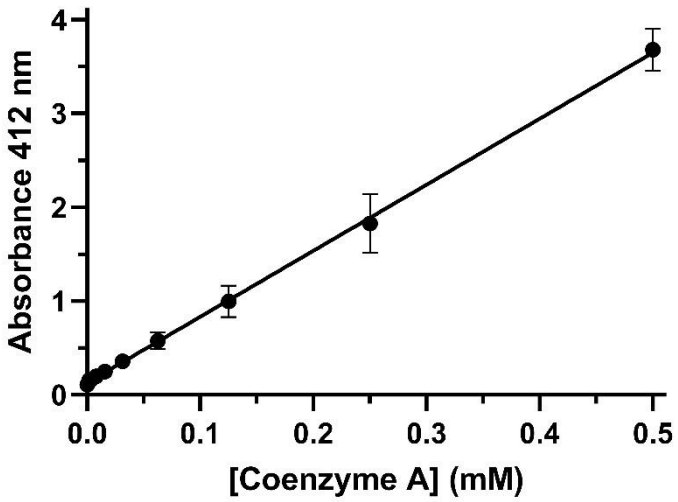


Figure B.2 DTNB Co-enzyme A standard curve. Co-enzyme A concentrations 2, 1, 0.5, 0.25 and 0.125 mM were collected. Final reaction volume of 250 µl per well. Datapoints are mean of triplicate values with SEM plotted as error bars.

Appendix C: Supplementary Material for Chapter Five

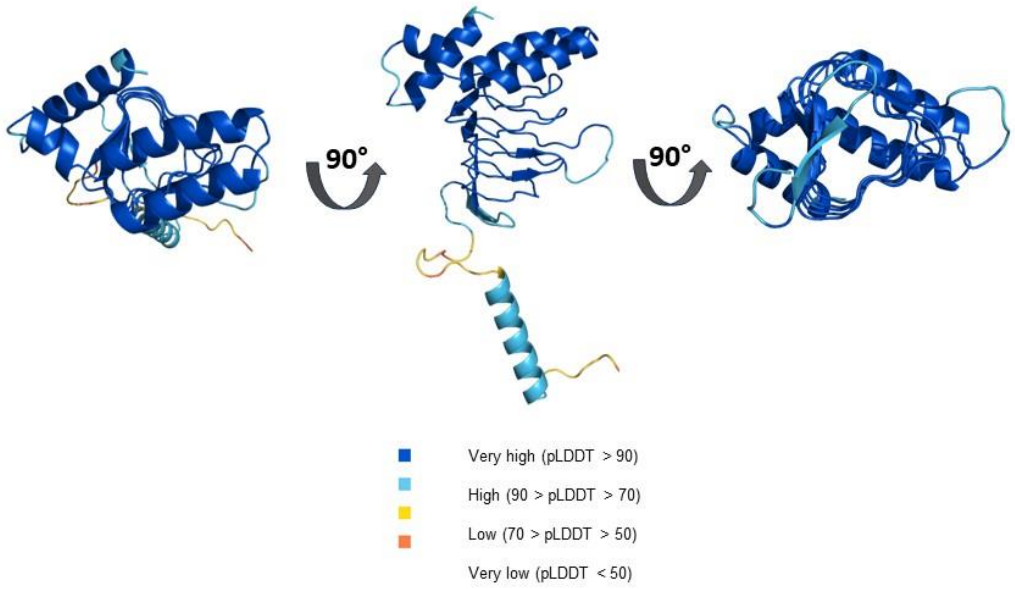


Figure C.1 Predicted *Sa*CysE monomer generated by ColabFold. Model is coloured by residue confidence scoring predicted local distance difference test (pLDDT); high confidence (blue, pLDDT > 90), confident (light blue, 90 > pLDDT > 70), low confidence (yellow, 70 > pLDDT > 50) and very low confidence (orange, pLDDT < 50). Protein sequence for *S. aureus* MW2 was retrieved from UniProt (P67766). The C-terminal tail is hidden in the most right figure for clarity. Figure created using PyMOL.

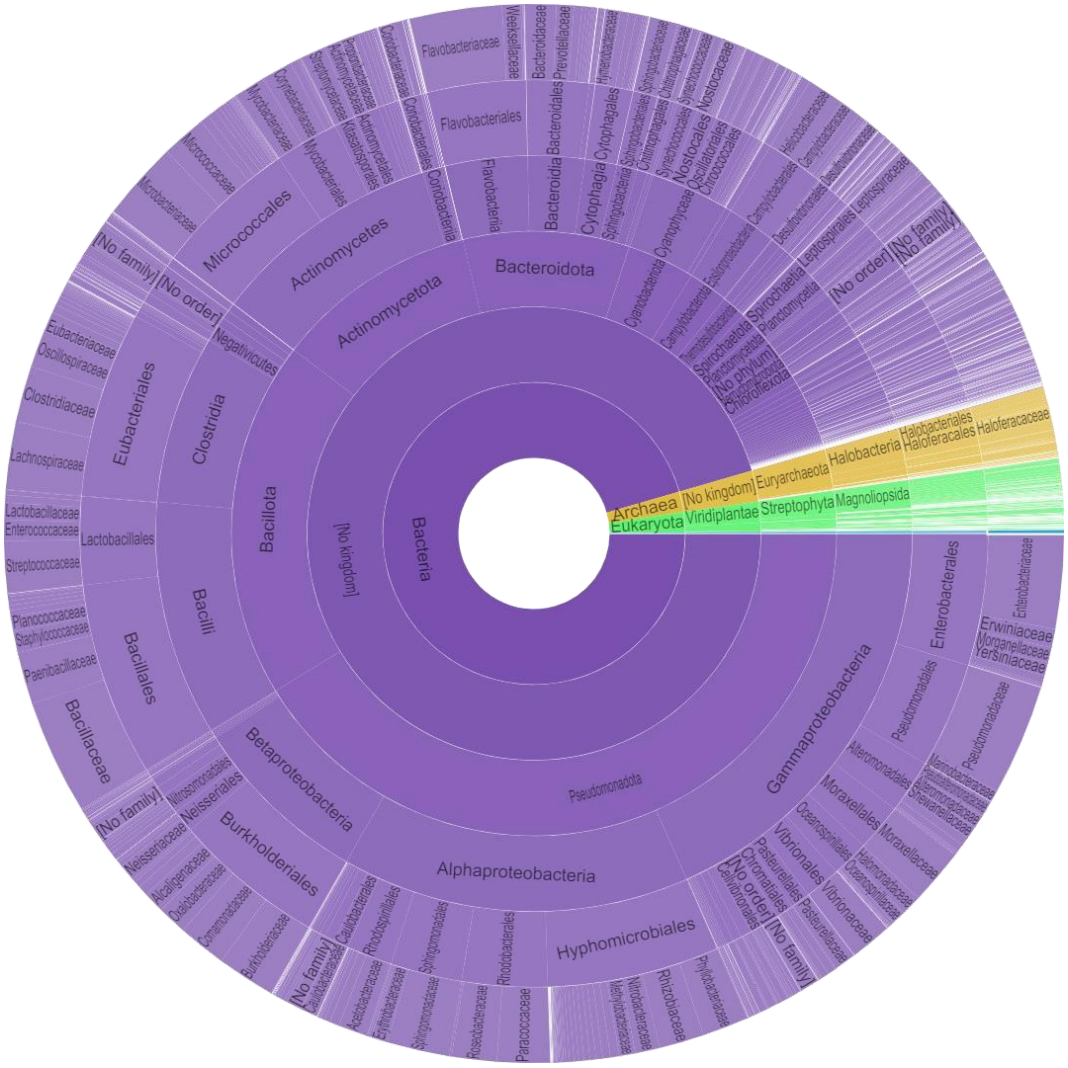


Figure C.2 Taxonomic distribution of species containing the SATase N-terminal domain (IPR042122). Sunburst diagram from InterPro (https://www.ebi.ac.uk/interpro/entry/InterPro) showing all species (21,995 species, 36,235 sequences) that are annotated with IPR042122. Sunburst coloured domain of life, with bacteria, archaea and eukaryota coloured, purple, orange, and green, respectively.

Table C.1 Published kinetic parameters for characterised CysE homologues.

Organism	K_m (mM) L-serine/acetyl-CoA	V_{max} (Mm.min^{-1}) serine/acetyl-CoA	K_{cat} (sec^{-1}) serine/acetyl-CoA	Specific activity (µmol.min^{-1}.mg^{-1})	N-terminal domain length IPR042122	Reference
Escherichia coli	1.3 ± 0.2/0.3 ± 0.1	-	323 ± 33	-	Full-length	(Benoni et al., 2017)
Haemophilus influenzae	1.17/0.2*	-	427	ND	Full-length	(Hindson & Shaw, 2003)
Salmonella typhimurium	5.6/1.1*	ND	ND	ND	Full-length	(Kredich & Tomkins, 1966)
Mycobacterium tuberculosis	0.0264 ± 0.0006/0.0513 ± 0.0050	0.0073 ± 0.0005	81.36 ± 5.22	10.66 ± 0.44	Truncated	(Qiu et al., 2013)
	2.1±0.3/ND#	ND	14.2 ± 0.3	ND		(Rahisuddin et al., 2024)
Mycolicibacterium smegmatis	ND	ND	N.D	0.0773 ± 0.0009	Truncated	(Qiu et al., 2014)
Staphylococcus aureus	ND	ND	ND	3,809.8 ± 24.3	Truncated	(Chen et al., 2019)
Neisseria gonorrhoeae	1.21± 0.16/ 0.15 ± 0.05	ND	1444 ± 41 / 1176 ± 111	ND	Full-length	(Oldham et al., 2022)
Brucella abortus	0.1275 ± 0.0199/ 0.1047±0.0125	ND	ND	ND	Full-length	(Kumar, Kumar, Alam, & Gourinath, 2014)

* No error reported
#trimeric
ND Not Determined

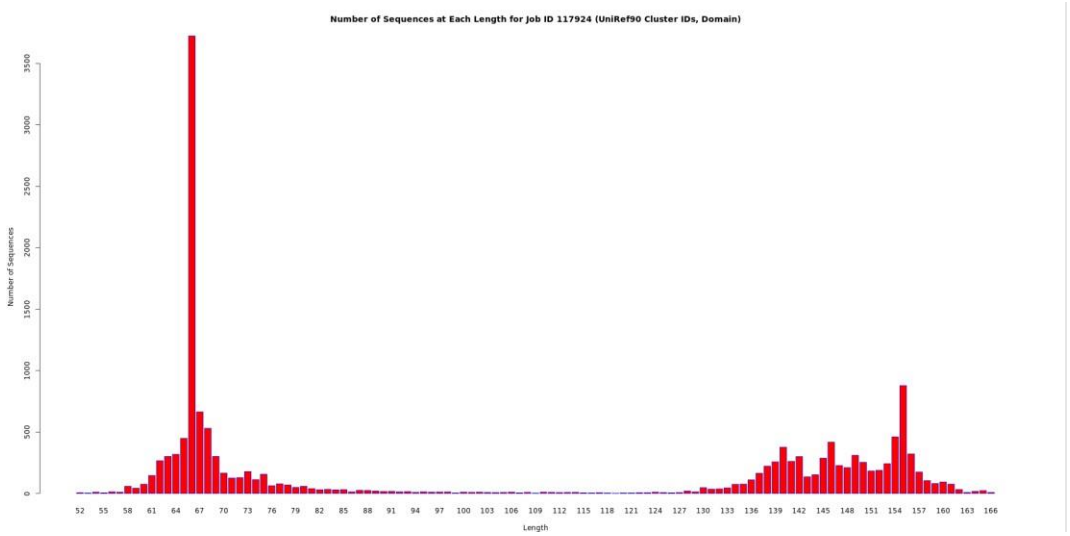


Figure C.3 IPR042122 domain length histogram of UniRef90 clusters. Figure produced by sequence similarity network tool using the online EFI-EST server (https://efi.igb.illinois.edu/efi-est/) (Oberg et al., 2023; Zallot et al., 2019).

Table C.2 Features of sequence similarity networks constructed with CysE SATase domain IPR042122.

Network name	Edge cutoff (% identity)	Cluster details	Domain length (residues)	Number of repnodes	Number of edges	Repnode cluster (%)
All CysE	20.8	Cluster I (truncated and full-length)	62-166	3914	154029	60
		Cluster II (full-length)	62-166	1181	523215	60
Truncated isoforms	33	Cluster A	62-77	3963	1207137	75
Cluster A	68.06	Cluster 1	62-77	2475	16335	75
		Cluster 2	62-77	56	740	75
		Cluster 3	62-77	66	213	75
Cluster 1	72	Cluster 1A	62-77	330	971	75
		Cluster 1B	62-77	300	985	75
		Cluster 1C	62-77	113	673	75
		Cluster 1D	62-77	268	1270	75

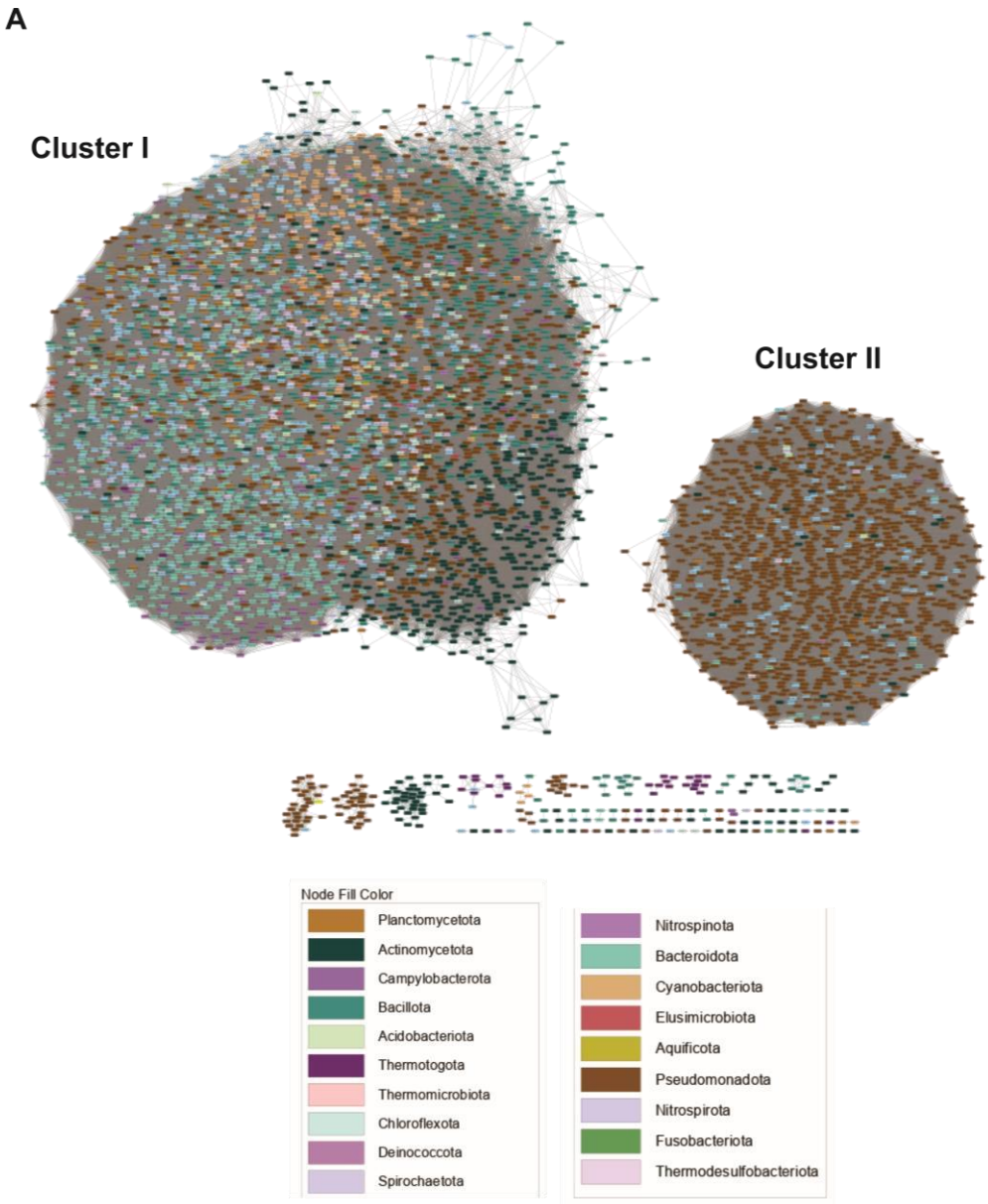


Figure C.4 Repnode CysE SSN coloured by phylum. Network is refined to an edge similarity cutoff of 20.8%. Nodes that belong to candidate phyla are coloured blue.

Table C.3 Details of bacterial CysE crystal structures.

Organism	Length	UniProt	Resolution (Å)	PDB	Reference
Escherichia coli	273	P0A9D4	2.2	1T3D	(Pye et al., 2004)
Haemophilus influenzae	267	P43886	2.7	1S80	(Gorman & Shapiro, 2004)
Phocaeicola vulgatus ATCC 8482	302	A6KZB9	2.00	3F1X	(Kuzin et al., 2008)
Yersinia pestis	273	A0A2U2H3H7	2.4	3GVD	(Kim, Zhou, Peterson, Anderson, & Joachimiak, 2006)
***Brucella abortus* 2308**	266	Q2YQG6	1.95	3MC4	(Seattle Structural Genomics Center for Infectious Disease (SSGCID), Abendroth, Gardberg, & Staker, 2010)
***Vibrio cholerae* O1 biovar El Tor str. N16961**	273	Q9KNT2	2.17 Å	4H7O	(Tarique, Abdul Rehman, & Gourinath, 2012)
***Brucella abortus* S19**	274	A0A0F6AR69	1.97	4HZC	(Kumar et al., 2014)
Planctopirus limnophila DSM 3776	311	D5SUT9	2.15	6LCN	(Kumar, Singh, & Kumaran, 2019)
***Klebsiella pneumoniae* MGH 78578**	273	A6TFK2	2.8	6JVU	(Verma et al., 2020)
Neisseria gonorrhoeae	272	Q5F6X0	2.01	6WYE	(Oldham et al., 2022)
Salmonella enterica subsp. enterica serovar Typhimurium	273	A0A0D6I3Y9	2.3	8I04	(Toyomoto et al., 2023)

Table C.4 List of sequence details used for iTOL tree annotation. Domain lengths are for N-terminal SATase domain IPR042122.

Species	Copies of truncated	Copies of full-length	Domain start	Domain end	Domain length	CysE protein length (AA)	Taxid
Thermoanaerobacter tengcongensis	1	0	1	66	66	220	273068
Clostridium acetobutylicum	1	0	1	66	66	186	272562
Clostridium perfringens	1	0	1	66	66	169	195103
Clostridium tetani	1	0	1	66	66	193	1513
Enterococcus faecalis	1	0	2	65	64	179	226185
Lactococcus lactis	1	0	9	70	62	199	1358
***Streptococcus pneumoniae* TIGR4**	1	0	3	65	63	205	170187
***Streptococcus pneumoniae* R6**	1	0	3	65	63	205	171101
Streptococcus mutans	1	0	15	82	68	222	210007
***Streptococcus agalactiae* V**	1	0	3	65	63	194	208435
***Streptococcus agalactiae* III**	1	0	3	65	63	194	211110
***Streptococcus pyogenes* M1**	1	0	4	65	62	193	310447
***Streptococcus pyogenes* MGAS8232**	1	0	4	65	62	193	186103
***Streptococcus pyogenes* SSI-1**	1	0	4	65	62	193	193567
***Streptococcus pyogenes* MGAS315**	1	0	4	65	62	193	198466
Staphylococcus epidermidis	1	0	1	66	66	213	176280
***Staphylococcus aureus* MW2**	1	0	1	66	66	213	196620
***Staphylococcus aureus* Mu50**	1	0	1	66	66	213	158878
***Staphylococcus aureus* N315**	1	0	1	66	66	213	158879
Listeria innocua	1	0	1	66	66	204	272626
***Listeria monocytogenes* EGD**	1	0	1	66	66	204	169963

***Listeria monocytogenes* F2365**	1	0	1	66	66	204	265669
Oceanobacillus iheyensis	1	0	1	68	68	226	221109
Bacillus halodurans	1	0	1	69	69	229	272558
Bacillus subtilis	1	0	2	67	66	217	224308
Bacillus anthracis	1	0	1	66	66	221	1392
***Bacillus cereus* ATCC 10987**	1	0	1	66	66	221	222523
***Bacillus cereus* ATCC 14579**	1	0	1	66	66	221	226900
Fibrobacter succinogenes	0	2	2	139	138	262	59374
Chlorobium tepidum	0	1	1	141	141	273	19439
Bacteroides thetaiotaomicron	3	1	34	190	157	300	226186
Rhodopirellula baltica	1	0	34	193	160	324	243090
Gemmata obscuriglobus	1	0	34	187	154	315	114
***Leptospira interrogans* 56601**	1	0	1	62	62	235	189518
***Leptospira interrogans* L1-130**	1	0	1	62	62	0	267671
Corynebacterium diphtheriae	1	0	1	69	69	188	1717
Corynebacterium efficiens	1	0	3	69	67	188	152794
Corynebacterium glutamicum	1	0	1	63	63	182	1718
***Corynebacterium glutamicum* 13032**	1	0	1	63	63	182	196627
Mycobacterium paratuberculosis	1	0	1	66	66	229	1770
Mycobacterium leprae	1	0	1	66	66	227	1769
Mycobacterium bovis	1	0	1	66	66	229	1765
***Mycobacterium tuberculosis* CDC1551**	1	0	1	66	66	229	911237
***Mycobacterium tuberculosis* H37Rv**	1	0	1	66	66	229	1773
Thermotoga maritima	2	0	4	80	77	220	243274
Dehalococcoides ethenogenes	1	0	1	62	62	230	243164

Gloeobacter violaceus	2	0	1	73	73	244	251221
Synechococcus elongatus	1	1	1	66	66	244	1140
***Nostoc* sp. PCC 7120**	2	0	1	66	66	253	103690
***Synechocystis* sp. PCC6803**	1	0	1	66	66	249	1111708
***Prochlorococcus marinus* CCMP1378**	1	0	1	66	66	244	59919
***Prochlorococcus marinus* SS120**	1	0	1	66	66	248	167539
***Prochlorococcus marinus* MIT9313**	1	0	1	71	71	251	74547
***Synechococcus* sp. WH8102**	1	0	1	68	68	248	84588
Solibacter usitatus	1	0	1	66	66	218	234267
Acidobacterium capsulatum	1	0	1	65	65	247	240015
Desulfovibrio vulgaris	0	1	42	196	155	323	882
Geobacter sulfurreducens	2	0	1	66	66	225	35554
Bdellovibrio bacteriovorus	1	0	5	66	62	168	264462
Campylobacter jejuni	1	0	3	68	66	212	192222
Wolinella succinogenes	1	0	1	72	72	237	273121
Helicobacter hepaticus	1	0	2	68	67	234	235279
***Helicobacter pylori* 26695**	1	0	3	65	63	171	85962
***Helicobacter pylori* J99**	1	0	3	65	63	171	85693
Caulobacter crescentus	0	1	1	149	149	279	190650
Rhodopseudomonas palustris	0	2	3	152	150	275	258594
Bradyrhizobium japonicum	0	1	5	151	147	274	375
Rhizobium loti	0	1	5	151	147	282	381
Brucella melitensis	0	1	1	152	152	274	546272
Brucella suis	0	1	1	152	152	274	470173
Rhizobium meliloti	0	1	2	152	151	275	266834

***Agrobacterium tumefaciens* WashU**	0	1	2	152	151	274	358
***Agrobacterium tumefaciens* Cereon**	0	1	2	152	151	274	358
Nitrosomonas europaea	0	1	44	197	154	308	228410
Chromobacterium violaceum	0	1	1	141	141	260	243365
***Neisseria meningitidis* B**	0	1	2	146	145	272	122586
***Neisseria meningitidis* A**	0	1	2	146	145	272	122587
Ralstonia solanacearum	0	1	1	66	66	248	1031711
Bordetella pertussis	0	1	44	198	155	317	257313
Bordetella bronchiseptica	0	1	44	198	155	317	257310
Bordetella parapertussis	0	1	44	198	155	317	257311
Pseudomonas aeruginosa	1	0	1	66	66	258	208964
Pseudomonas putida	1	0	1	66	66	261	303
Pseudomonas syringae	0	1	1	135	135	254	317
Shewanella oneidensis	1	0	2	68	67	177	211586
Photobacterium profundum	0	1	1	142	142	273	298386
Vibrio cholerae	0	1	1	142	142	273	345073
Vibrio parahaemolyticus	0	1	1	142	142	273	670
***Vibrio vulnificus* YJ016**	0	1	1	142	142	273	196600
***Vibrio vulnificus* CMCP6**	0	1	1	142	142	273	216895
Haemophilus ducreyi	0	1	1	142	142	271	233412
Haemophilus influenzae	0	1	1	138	138	267	71421
Pasteurella multocida	0	1	3	138	136	264	272843
Blochmannia floridanus	0	1	1	141	141	244	203907
***Buchnera aphidicola* Bp**	1	1	1	143	143	254	224915

***Buchnera aphidicola* APS**	0	1	4	142	139	274	107806
***Buchnera aphidicola* Sg**	0	1	3	142	140	261	198804
Photorhabdus luminescens	0	1	1	142	142	273	29488
***Yersinia pestis* Medievalis**	1	0	1	142	142	273	1234662
***Yersinia pestis* KIM**	0	1	1	142	142	273	187410
***Yersinia pestis* CO92**	0	1	1	142	142	273	214092
Salmonella typhimurium	0	1	1	142	142	273	99287
Salmonella enterica	0	1	1	142	142	273	28901
Salmonella typhi	0	1	1	142	142	273	90370
***Escherichia coli* EDL933**	0	1	1	142	142	273	155864
***Escherichia coli* O157 H7**	0	1	1	142	142	273	83334
***Shigella flexneri* 2a 2457T**	0	1	1	142	142	273	198215
***Shigella flexneri* 2a 301**	0	1	1	142	142	273	198214
***Escherichia coli* O6**	0	1	1	142	142	273	1370107
***Escherichia coli* K12**	0	1	1	142	142	273	83333

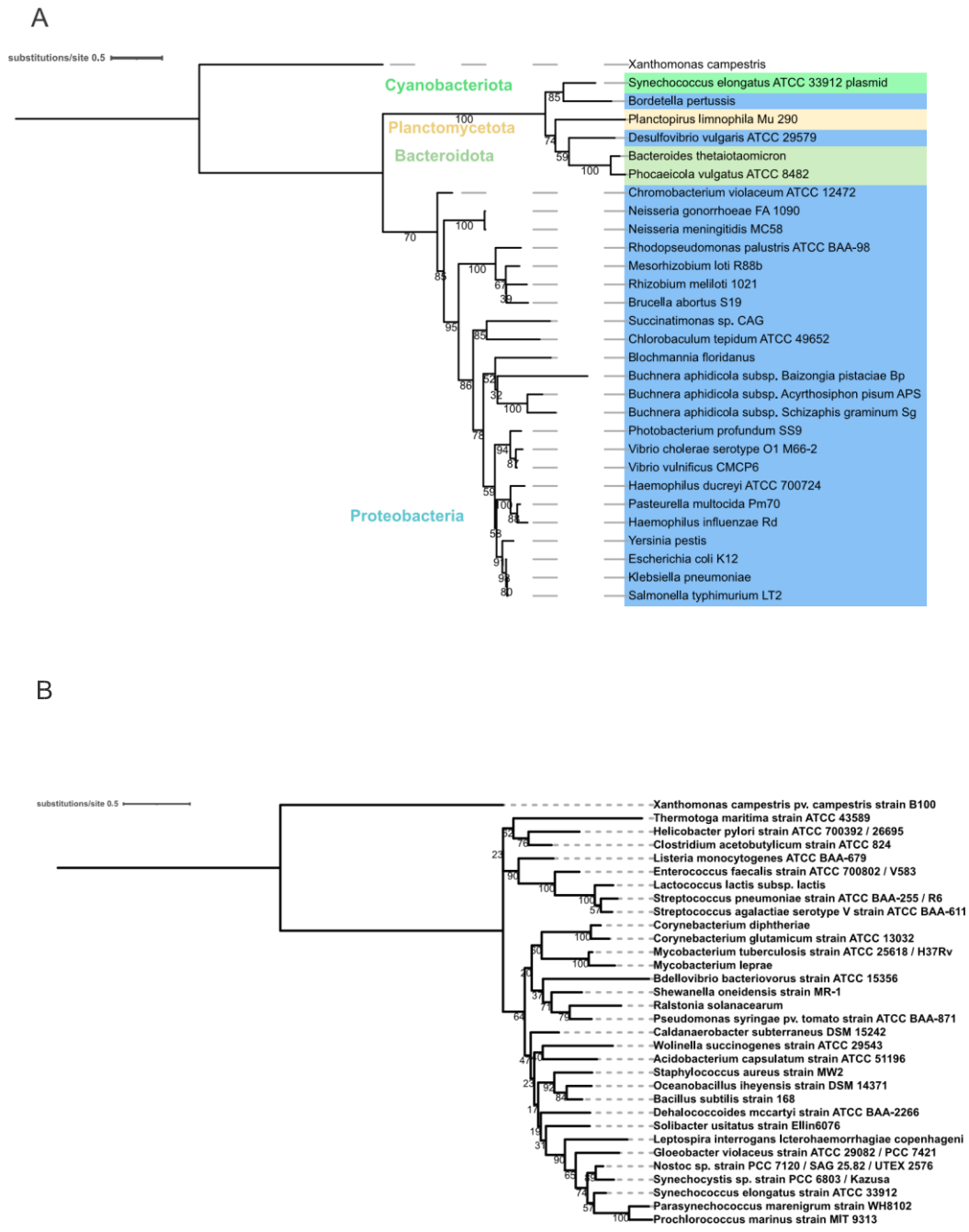


Figure C.5 Maximum-likelihood trees of full-length and truncated CysE isoform sequences. (A) Maximum-likelihood tree of full-length CysE isoforms. (B) Maximum-likelihood tree of truncated CysE isoforms. CysE protein sequences were retrieved for iTOL bacterial species and were categorised as either full-length or truncated based on the length of the N-terminal SATase domain. Trees were constructed using a trimmed alignment of both the N-terminal SATase domain and left-handed β-helix domain. All sequences were chromosomal encoded with the exception of the full-length

Synechococcus elongatus which is plasmid encoded. Maltose *O*-acetyltransferase (LβH member) from *Xanthomonas campestris* (UniProt B0RYQ0) was used as an outgroup. Branch labels are supporting bootstrap confidence scores. Bootstraps were calculated using ultrafast bootstrap (UFBoot) (Hoang, Chernomor, von Haeseler, Minh, & Vinh, 2017). Figure was created using iTOL tree viewer (Letunic & Bork, 2021).

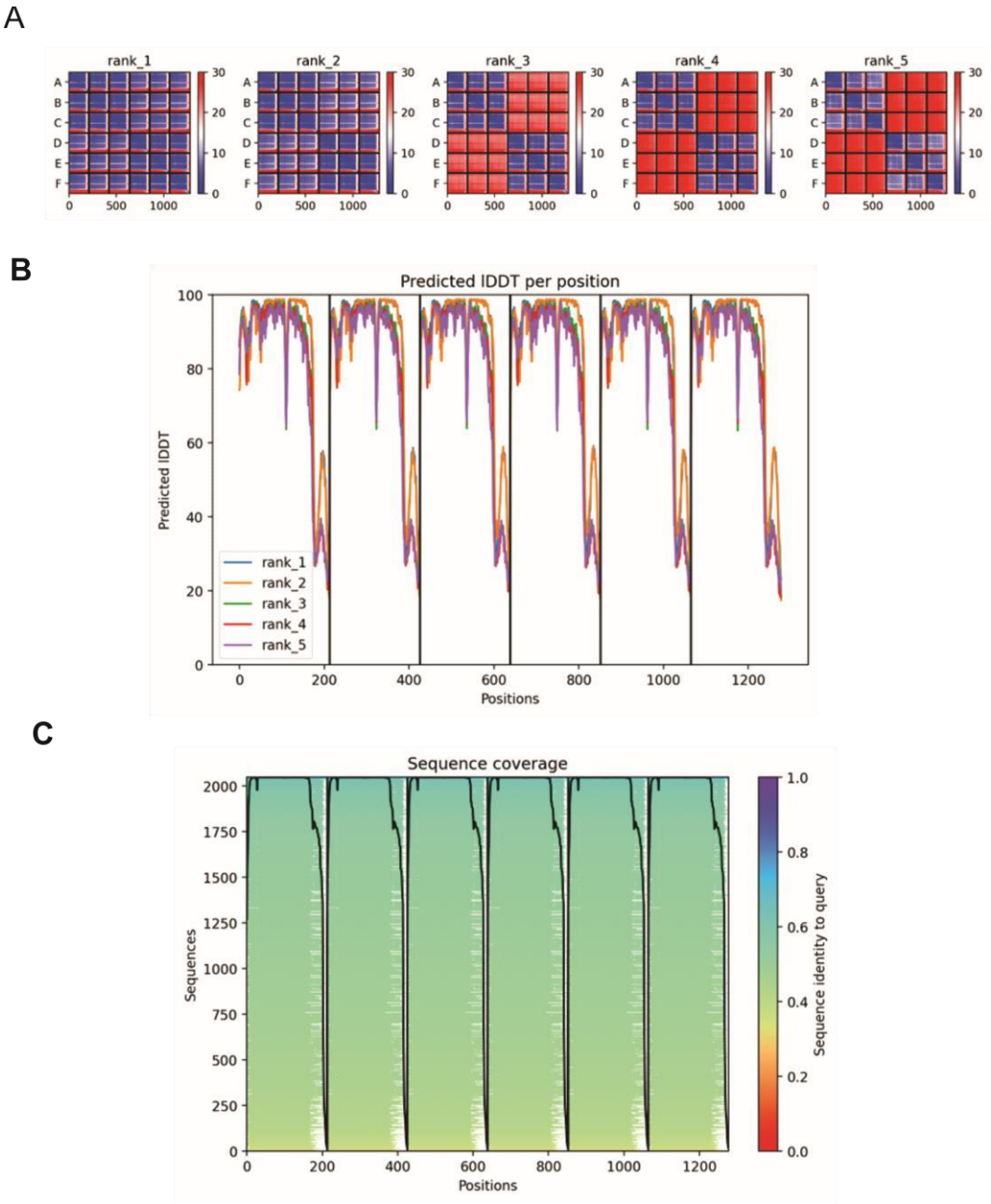


Figure C.6 ColabFold scores and statistics for *Sa*CysE model. ColabFold models and statistics for model generated for *Sa*CysE. (A) predicted error alignment for each of the five ranked models, (C) predicted local distance difference test (pLDDT) scoring per model, and (D) sequence coverage per model.

Appendix D: Co-Authorship forms

D.1 Co-authorship form for Chapter Two: Literature Review

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

Postgraduate Studies Office
Student and Academic Services Division
Wahanga Ratonga Matauranga Akonga
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 4439
Website: http://www.waikato.ac.nz/sasd/postgraduate/

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 2: literature review

Hicks, J. L., Oldham, K. E. A., McGarvie, J., & Walker, E. J. (2022). Combatting antimicrobial resistance via the cysteine biosynthesis pathway in bacterial pathogens. Biosci Rep, 42(10) 10.1042/bsr20220368

Nature of contribution by PhD candidate	Analysis of literature, figure creation, drafting and editing of manuscript.
Extent of contribution by PhD candidate (%)	30

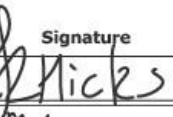
CO-AUTHORS

Name	Nature of Contribution
Joanna Hicks	Drafting and editing of manuscript, project administration and funding acquisition
Jack McGarvie	Analysis of literature, figure creation, drafting, and editing, of manuscript.
Emma Walker	Analysis of literature, figure creation, drafting and editing of manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and

Name	Signature	Date
Joanna Hicks		24/01/24
Jack McGarvie		08/02/224
Emma Walker		05/02/24

July 2015

D.2 Co-authorship form for Chapter Three: Serine acetyltransferase from *Neisseria gonorrhoeae*; structural and biochemical basis of inhibition.

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

Postgraduate Studies Office
Student and Academic Services Division
Wahanga Ratonga Matauranga Akonga
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 4439
Website: http://www.waikato.ac.nz/sasd/postgraduate/

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 3:
Oldham, K. E. A., Prentice, E. J., Summers, E. L., & Hicks, J. L. (2022). Serine acetyltransferase from Neisseria gonorrhoeae; structural and biochemical basis of inhibition. The Biochemical journal, 479(1), 57-74. doi:10.1042/bcj20210564 (Published).

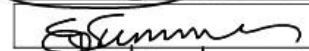
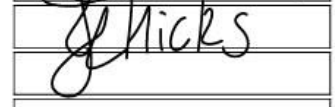
Nature of contribution by PhD candidate	data collection, processing and analysis; Method optimization; drafting and editing of manuscript.
Extent of contribution by PhD candidate (%)	85

CO-AUTHORS

Name	Nature of Contribution
Erica Prentice	Analysis and interpretation of kinetic data; drafting and editing of manuscript.
Emma Summers	Analysis and interpretation of protein structure, drafting and editing of manuscript.
Joanna Hicks	Data collection, project supervision, drafting and editing of manuscript

Certification by Co-Authors

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and

Name	Signature	Date
Erica Prentice		6/2/24
Emma Summers		7/2/2024
Joanna Hicks		24/01/24

July 2015

D3. Co-authorship form for Chapter Four: Identification of novel inhibitors of serine acetyltransferase from *Neisseria gonorrhoeae*

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

Postgraduate Studies Office
Student and Academic Services Division
Wahanga Ratonga Matauranga Akonga
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 4439
Website: http://www.waikato.ac.nz/sasd/postgraduate/

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.
Chapter 4: identification of novel inhibitors of serine acetyltransferase from Neisseria gonorrhoeae.

Oldham, K. E. A., Jiao, W., & Hicks, J. L. (2023). Identification of novel inhibitors targeting serine acetyltransferase from Neisseria gonorrhoeae. (in preparation).

Nature of contribution by PhD candidate	Data collection and analysis, method optimization, figure creation, drafting and editing of manuscript.
Extent of contribution by PhD candidate (%)	80

CO-AUTHORS

Name	Nature of Contribution
Wanting Jiao	Data collection and analysis, figure creation, drafting and editing of manuscript
Joanna Hicks	Drafting and editing of manuscript, project administration and funding acquisition

Certification by Co-Authors

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and

Name	Signature	Date
Wanting Jiao	Wanting Jiao	12/02/24
Joanna Hicks	J Hicks	24/01/24

July 2015

D4. Co-authorship form for Chapter Five: Curious case of CysE: Diversity and Distribution of Serine Acetyltransferases

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

Postgraduate Studies Office
Student and Academic Services Division
Wahanga Ratonga Matauranga Akonga
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 4439
Website: http://www.waikato.ac.nz/sasd/postgraduate/

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. Chapter five: The curious case of CysE: Diversity and distribution of truncated serine O-acetyltransferases Oldham, K. E. A., Grout, E. K., Williamson, A., & Hicks, J. L. The curious case of CysE: diversity and distribution of truncated serine O-acetyltransferases.(Prepared for submission)

Nature of contribution by PhD candidate	Data analysis, figure preparation, manuscript drafting and editing.
Extent of contribution by PhD candidate (%)	90

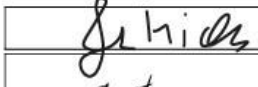
CO-AUTHORS

Name	Nature of Contribution
Joanna Hicks	Project supervision, funding acquisition, manuscript drafting and editing
Adele Williamson	data analysis and manuscript editing.
Emily Grout	figure preparation, data analysis and manuscript editing.

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and

Name	Signature	Date
Joanna Hicks		26/02/24
Adele Williamson		29/02/2024
Emily Grout		26/02/24

July 2015