



MONASH University

Inhibitors of the folate pathway enzymes HPPK and DHPS

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

By

Matthew L. Dennis

December 2016

Department of Medicinal Chemistry

Monash Institute of Pharmaceutical Sciences (MIPS)

Monash University

Copyright Notices

Notice 1

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

© Matthew L. Dennis (2016)

Notice 2

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

© Matthew L. Dennis (2016)

Abstract

Antibiotic resistance represents an increasing worldwide healthcare concern. The pursuit of novel and effective antibiotics to overcome this resistance is of grave necessity. Sadly, the challenges and high costs associated with antibiotic development represent a significant deterrent for pharmaceutical companies, who are able to generate much higher profits from drugs that are taken on a regular basis.

Folate is an essential vitamin (vitamin B9), necessary for various one-carbon transfer reactions and essential for the synthesis of a range of biomolecules, including thymidine, glycine, and methionine. While humans depend on both passive and active mechanisms for uptake of exogenous folate, bacteria synthesise folate *de novo*. Many enzymes of the bacterial folate biosynthesis pathway are absent from humans, and therefore they represent attractive drug targets. Indeed, the folate pathway enzyme dihydropteroate synthase (DHPS) is the target of the long-standing 'sulfa' class of antibiotics. These drugs have enjoyed success in treatment against various infectious organisms with sulfamethoxazole (in combination with trimethoprim) remaining the gold standard for prophylaxis and treatment of *Pneumocystis jirovecii* pneumonia (the most common fungal infection in HIV and AIDS infected patients). In general, however, the utility and therapeutic impact of these drugs has diminished over the years due to antibiotic resistance. Novel antimicrobials that can combat resistant pathogens are required, and enzymes of the folate pathway remain promising targets.

6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) is the second enzyme of the folate biosynthetic pathway and is currently not the known target of any drug. HPPK catalyses the pyrophosphate transfer from an adenosine triphosphate (ATP) molecule to the pterin substrate, generating dihydropterin pyrophosphate (PtPP). The adjacent downstream enzyme, DHPS, synthesises dihydropteroate from PtPP and a *para*-aminobenzoic acid (pABA) molecule. Sulfa drugs bind in the pABA-binding pocket and competitively inhibit this enzyme, in addition to acting as "dead-end" substrates that may both deplete cellular PtPP and inhibit downstream enzymes. Mutations in DHPS that have conferred antibiotic resistance are predominantly observed within the loop regions around the pABA-binding pocket. The rigid pterin-binding pocket, located deep within the DHPS enzyme, is well conserved between species. This rigidity and substrate selectivity is

anticipated to reduce the likelihood of catalytically competent resistant strains emerging through mutations within the pterin site.

The work presented in this thesis represents a continuation of previous work carried out within our group, which identified a hit fragment, 8-mercaptoguanine (8MG), that binds to the pterin-site of HPPK from *Staphylococcus aureus* (SaHPPK). Extension from the thiol group of 8MG appeared a promising strategy and one commercially available derivative, 8-((2-(4-methoxyphenyl)-2-oxoethyl)thio)guanine (compound **3** in the first publication presented), was shown to bind SaHPPK with low micromolar affinity.

The contents of the three papers that form the main body of this thesis are as follows. Paper 1: the SaHPPK binding affinity of **3** is quantified and the binding mode determined, through both NMR studies and a crystal structure of **3** in complex with cofactor-bound SaHPPK. A series of synthesised analogues of **3** are tested — a subset of these are presented in this paper — leading to identification of a new lead compound that improves upon both the affinity and physicochemical properties of **3**. An additional ternary ligand/cofactor/SaHPPK crystal structure is determined, as well as the binary cofactor/SaHPPK complex, providing insight into the nature of the protein-ligand interactions and the large protein conformational changes that occur upon ligand binding.

Paper 2: Further synthesised 8MG derivatives are tested to probe the structure-activity relationship (SAR) for SaHPPK binding. Compounds are also tested in parallel against HPPK from the Gram-negative bacterium, *Escherichia coli* (EcHPPK). Clear species selectivity is observed for an 8MG-benzyl scaffold, while no selectivity is observed for an 8MG-acetophenone scaffold. This is rationalised via 11 new crystal structures of compounds bound to both SaHPPK and EcHPPK that reveal the formation of a novel binding pocket, one that is both species and compound dependent. The top-binding compound displayed a binding constant of 120 nM for SaHPPK binding, the highest reported affinity for an HPPK inhibitor to date, with a selectivity index of ~15 between the two species.

Paper 3: The synthesised 8MG-derivatives are tested against DHPS from *E. coli* (EcDHPS). Crystal structures of various lead compounds guide the development of novel inhibitors and the resulting compounds display up to 100-fold greater affinity than the parent, 8MG. In conjunction with membrane-permeabilising agents, antibacterial activity is observed for many of the compounds tested. A crystal structure of EcDHPS in complex with its catalysis product is additionally obtained.

Overall, the current study describes the significant advancement of the 8MG compound series as inhibitors of the antibiotic folate pathway targets HPPK and DHPS. Greater than 100-fold improvements in affinity for both proteins have been obtained, relative to the parent compound, and rationalised on the basis of structural data. The tightest binding compounds and the corresponding structural data constitute a significant step forward in terms of the design of new compounds that may contribute to the fight against antibiotic resistance.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer reviewed journals and one prepared manuscript. The core theme of the thesis is the development of inhibitors of the enzyme SaHPPK, an antibacterial target, with a focus on structurally-determined binding modes. The ideas, development and drafting of all of the papers in the thesis were the principal responsibility of myself, the student, working within the Medicinal Chemistry Theme under the supervision of Dr James D. Swarbrick, Dr Thomas Peat (CSIRO), Prof. Jonathan B. Baell, and Assoc. Prof. Bim Graham.

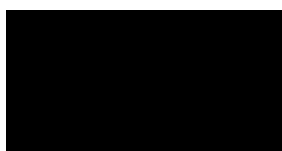
The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of publications, my experimental contribution to the work involved the following: all biochemical and biophysical methods employed (from plasmid to PDB) including protein expression and purification, as well as NMR spectroscopy, X-ray crystallography, surface plasmon resonance, isothermal titration calorimetry, and enzymatic assays.

Assistance (particularly at the start of this thesis) in the above techniques was provided by: Dr Sandeep Chhabra – protein expression and purification as well as enzymatic assays; Dr Tom Peat – X-ray crystallography; Dr James Swarbrick – NMR spectroscopy; Dr Olan Dolezal – Surface Plasmon Resonance; Meghan Hattarki – Surface Plasmon Resonance and Isothermal Titration Calorimetry; Janet Newman – crystallisation and differential scanning fluorimetry. The liquid handling robotics at the Collaborative Crystallisation Centre (C3), responsible for setting up crystallisation and differential scanning fluorimetry plates, were operated by Dr Janet Newman. The synthetic chemistry presented in this thesis was performed by Dr Zhong-Chang Wang, Dr Aaron DeBono, Noel Pitcher, Dr Jitendra Harjani, and Dr Michael Lee.

Thesis Chapter	Publication Title	Publication status	Nature and extent of candidate's contribution
2	Structure-based design and development of functionalized mercaptoguanine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from <i>Staphylococcus aureus</i>	Published	Experimental design. Conduct of laboratory work. Data analysis. Preparation of manuscripts and subsequent revisions. Total contribution 70%
3	Structural basis for the selective binding of inhibitors to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from <i>Staphylococcus aureus</i> and <i>Escherichia coli</i>	Published	Experimental design. Conduct of laboratory work. Data analysis. Preparation of manuscripts and subsequent revisions. Total contribution 70%
4	8-Mercaptoguanine derivatives as inhibitors of dihydropteroate synthase from <i>Escherichia coli</i>	To be submitted	Experimental design. Conduct of laboratory work. Data analysis. Preparation of manuscripts and subsequent revisions. Total contribution 70%

Student signature:



Date: 19/12/2016

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature:



Date: 19/12/2016

Publication Record:**Publications during enrolment:**

1. **Dennis, M. L.**; Chhabra, S.; Wang, Z.-C.; DeBono, A. J.; Dolezal, O.; Newman, J.; Pitcher, N. P.; Rahmani, R.; Cleary, B.; Barlow, N.; Hattarki, M.; Graham, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structure-based design and development of functionalized mercaptoguanine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*. *Journal of Medicinal Chemistry* **2014**, 57, 9612–9626
2. **Dennis, M. L.**; Pitcher, N. P.; Lee, M. D.; DeBono, A. J.; Wang, Z.-C.; Harjani, J. R.; Rahmani, R.; Cleary, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structural basis for the selective binding of inhibitors to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus* and *Escherichia coli*. *Journal of Medicinal Chemistry* **2016**, 59, 5248–5263

Additional publications accepted during the thesis period:

1. Lee, M. D.; Loh, C. T.; Shin, J.; Chhabra, S.; **Dennis, M. L.**; Otting, G.; Swarbrick, J. D.; Graham, B. Compact, hydrophilic, lanthanide-binding tags for paramagnetic NMR spectroscopy. *Chemical Science* **2015**, 6, 2614–2624.
2. Swarbrick, J. D.; Ung, P.; **Dennis, M. L.**; Lee, M. D.; Chhabra, S.; Graham, B. Installation of a Rigid EDTA-Like Motif into a Protein α -Helix for Paramagnetic NMR Spectroscopy with Cobalt(II) Ions. *Chemistry – A European Journal* **2016**, 22, 1228–1232.
3. Lee, M. D.; **Dennis, M. L.**; Swarbrick, J. D.; Graham, B. Enantiomeric two-armed lanthanide-binding tags for complementary effects in paramagnetic NMR spectroscopy. *Chemical Communications* **2016**, 52, 7954–7957.

DEDICATION

To my parents.

Acknowledgements

I would like to express my gratitude to all those who have contributed to the realisation of my thesis. First and foremost, I would like to thank my research supervisor Dr James Swarbrick for the opportunity to work as his student. I would like to thank him for teaching me how to run NMR experiments and assign NMR spectra, in addition to all the various little things taught these last years. His continued guidance, support, and vision was fundamental to the success of this project. His efforts and attention to detail in both experimentation and writing were beyond any expectations, and have elevated this thesis above what it would have otherwise been.

I am greatly thankful to Dr Tom Peat for his mentorship throughout this project. His expertise in all things crystallographic and time spent teaching me has been invaluable. Discussions of refractory crystals and proteins helped immensely in generating new ideas that pushed this project forward. In particular his advice regarding subjects outside of this thesis, and to my future scientific endeavours, is very much appreciated.

I am thankful to Prof. Jonathan Baell for his large contribution. His management of both the medicinal and synthetic chemistry aspects of this project was outstanding and I am deeply indebted to both him and the chemists who performed the laboratory chemistry work: Dr Aaron DeBono, Noel Pitcher, Dr Jitendra Harjani, and Dr Zhong-Chang Wang as well as the supervisory roles of Dr Raphaël Rahmani and Dr Ben Cleary.

I would like to thank Assoc. Prof. Bim Graham, his calm pragmatism provided an appropriate balance to the optimistic enthusiasm of James. His suggestions and input on both the experimental and writing side are much appreciated.

I am grateful to Dr Sandeep Chhabra for helping me with protein expression and purification, as well as inhibition assay techniques. I am thankful to Dr Janet Newman for her help with crystallization techniques – her enthusiasm and expertise is second-to-none and I am lucky to have learnt from someone of her calibre. I would also like to acknowledge the help of Dr Shane Seabrook for his help with both crystallization and DSF experiments.

I am grateful to the teaching and assistance provided by Dr Olan Dolezal and Meghan Hattarki in SPR experimentation, and for the use of Olan's SPR equipment. Olan's insistence on my independence was exactly what I needed – the trust and responsibility

placed in me highly motivated me to improve my scientific abilities, and I am thankful to have had access to an environment that allows me to thrive in this manner.

I am greatly thankful to Dr Michael Lee for the synthetic work performed and his methodical attention to detail therewith. I am grateful to have had someone with whom I could share complaints about weekend NMR work and other scientific inconveniences.

I would like to thank Assoc. Prof. Martin Scanlon for the use of his lab facilities, as well as

Dr Jason Dang for his assistance with the NMR magnets. I am thankful to all my lab buddies and office buddies whose helpful discussions have helped me throughout this project, in particular Trayder Thomas for his help with molecular dynamics and Tamir Dingjan for his aesthetic eye and skills at creating pictures. I would also like to thank Dr Krithika Sundaram for her support throughout my PhD project.

I thank Monash Faculty of Pharmacy and Pharmaceutical Sciences for financial support during the course of my PhD. I thank CMSE (CSIRO, Parkville campus) for the top-up scholarship. I would also like to thank the C3 crystallization facility and Australian synchrotron for data collection. Thank you to all the staff at the MIPS and CSIRO for your support in all these years

I would like to thank Estelle Suys for her ongoing support, encouragement, and belief in me. Finally, I would like to thank my family without whom I would not have reached this milestone in my career.

Table of Contents

Abstract.....	I
Thesis including published works declaration.....	IV
Publication record.....	VII
Dedication.....	IX
Acknowledgements.....	X
Table of Contents.....	XII
List of Abbreviations.....	XIV
1. Introduction.....	1
1.1 Antibiotic resistance.....	2
1.2 <i>Staphylococcus aureus</i>	3
1.3 <i>Escherichia coli</i>	4
1.4 The folate pathway.....	5
1.5 HPPK.....	7
1.6 DHPS.....	13
1.7 Hit development in drug discovery.....	17
1.7.1 Fragment-based drug discovery.....	17
1.7.1.1 Structure-based drug design.....	19
1.8 Scope of the thesis.....	20
1.9 References.....	22
2. Structure-based design and development of functionalized mercaptoguanine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from <i>Staphylococcus aureus</i>	39
2.1 Preface to Chapter 2.....	40
2.2 Declaration for Thesis Chapter 2.....	43
2.3 Journal of Medicinal Chemistry (2014) publication.....	46
2.2 Supporting Information.....	61
3. Structural basis for the selective binding of inhibitors to 6- hydroxymethyl-7,8-dihydropterin pyrophosphokinase from <i>Staphylococcus aureus</i> and <i>Escherichia coli</i>	63

3.1 Preface to Chapter 3.....	64
3.2 Declaration for Thesis Chapter 3.....	68
3.3 Journal of Medicinal Chemistry (2016) publication.....	71
3.4 Supporting Information.....	87
4. 8-Mercaptoguanine derivatives as inhibitors of dihydropteroate synthase from <i>Escherichia coli</i>	113
4.1 Preface to Chapter 4.....	114
4.2 Declaration for Thesis Chapter 4.....	117
4.3 Prepared manuscript.....	120
4.3 Supporting Information.....	153
5. Materials and Methods.....	173
5.1 Introduction.....	174
5.2 Compound procurement and synthesis.....	174
5.3 X-ray crystallography.....	177
5.4 Bacterial cell preparation.....	187
5.5 Protein expression.....	187
5.6 Purification.....	190
5.7 Surface plasmon resonance.....	191
5.8 NMR spectroscopy.....	196
5.9 Molecular modelling.....	196
5.10 Isothermal titration calorimetry.....	197
5.11 Differential scanning fluorimetry.....	198
5.12 KinaseGlo biochemical assay.....	199
5.13 In vitro antibacterial assay.....	200
5.14 References.....	203
6. Conclusion.....	207
Appendix I – Media recipes.....	210
Appendix II – Additional publications during the thesis period.....	212
Permissions.....	233

List of Abbreviations

8MG – 8-mercaptopguanine

ACME – arginine catabolic mobile element

ADP – adenosine diphosphate

AMP – adenosine monophosphate

AMPCPP – α,β -methyleneadenosine 5'-triphosphate

ATP – adenosine triphosphate

ATPase – adenosine triphosphatase

BaDHPS – dihydropteroate synthase from *Bacillus anthracis*

BSA – bovine serum albumin

C3 – CSIRO Collaborative Crystallisation Centre

CIF – crystallographic information file

DHF – 7,8-dihydropteroyl glutamate

DHFS – dihydrofolate synthase

DHFR – dihydrofolate reductase

DHNA – dihydroneopterin aldolase

DHP – 7,8-dihydropteroate

DHPS – dihydropteroate synthase

DMSO – dimethyl sulfoxide

DNA – deoxyribonucleic acid

DNase – deoxyribonuclease

DTT – dithiothreitol

DSF – differential scanning fluorimetry

E. coli – *Escherichia coli*

EcHPPK – 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Escherichia coli*

EcDHPS – dihydropteroate synthase from *Escherichia coli*

EDC – *N*-ethyl-*N'*-(3-diethylaminopropyl)carbodiimide

EDTA – ethylenediaminetetraacetic acid

FBDD – fragment-based drug discovery

FDA – Food and Drug Administration

FPLC – fast protein liquid chromatography
GTP – guanosine triphosphate
GRP-78 – 78 kDa glucose-regulated protein
GTPCH – GTP cyclohydrolase
HBS – HEPES buffered saline
HEPES – 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HMDP – 7,8-dihydro-6-hydroxymethylpterin
HMDP-PP – 7,8-dihydro-6-hydroxymethylpterin pyrophosphate
HPPK – 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase
HTS – high-throughput screening
IMAC – immobilized metal affinity chromatography
ITC – isothermal titration calorimetry
IPTG – isopropyl β -D-1-thiogalactopyranoside
LB – lysogeny broth
MIC – minimum inhibitory concentration
MPD – 2-methyl-2,4-pentanediol
MPEG – methoxy poly(ethylene glycol)
MRSA – methicillin-resistant *Staphylococcus aureus*
MW – molecular weight
NHS – *N*-hydroxysuccinimide
NMR – nuclear magnetic resonance
NOE – nuclear Overhauser effect
NOESY – nuclear Overhauser effect spectroscopy
NTA – nitrilotriacetic acid
OD₆₀₀ – optical density at 600 nm
pABA – *para*-aminobenzoic acid
PDB – Protein Data Bank
PEG – polyethylene glycol
Pi – phosphate
RMSD – root-mean-square deviation
rpm – revolutions per minute

S. aureus – *Staphylococcus aureus*

SaHPPK – 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*

SAR – structure-activity relationship

SBDD – structure-based drug design

SDS – sodium dodecyl sulfate

SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEM – standard error of the mean

SER – surface entropy reduction

SPR – surface plasmon resonance

Sulfa – sulfonamide class of antibiotics

TEMED – tetramethylethylenediamine

TEV – tobacco etch virus protease

TIM-barrel – triosephosphate isomerase barrel

THF – tetrahydrofolate

Tris – tris(hydroxymethyl)aminomethane

TROSY – transverse relaxation-optimized spectroscopy

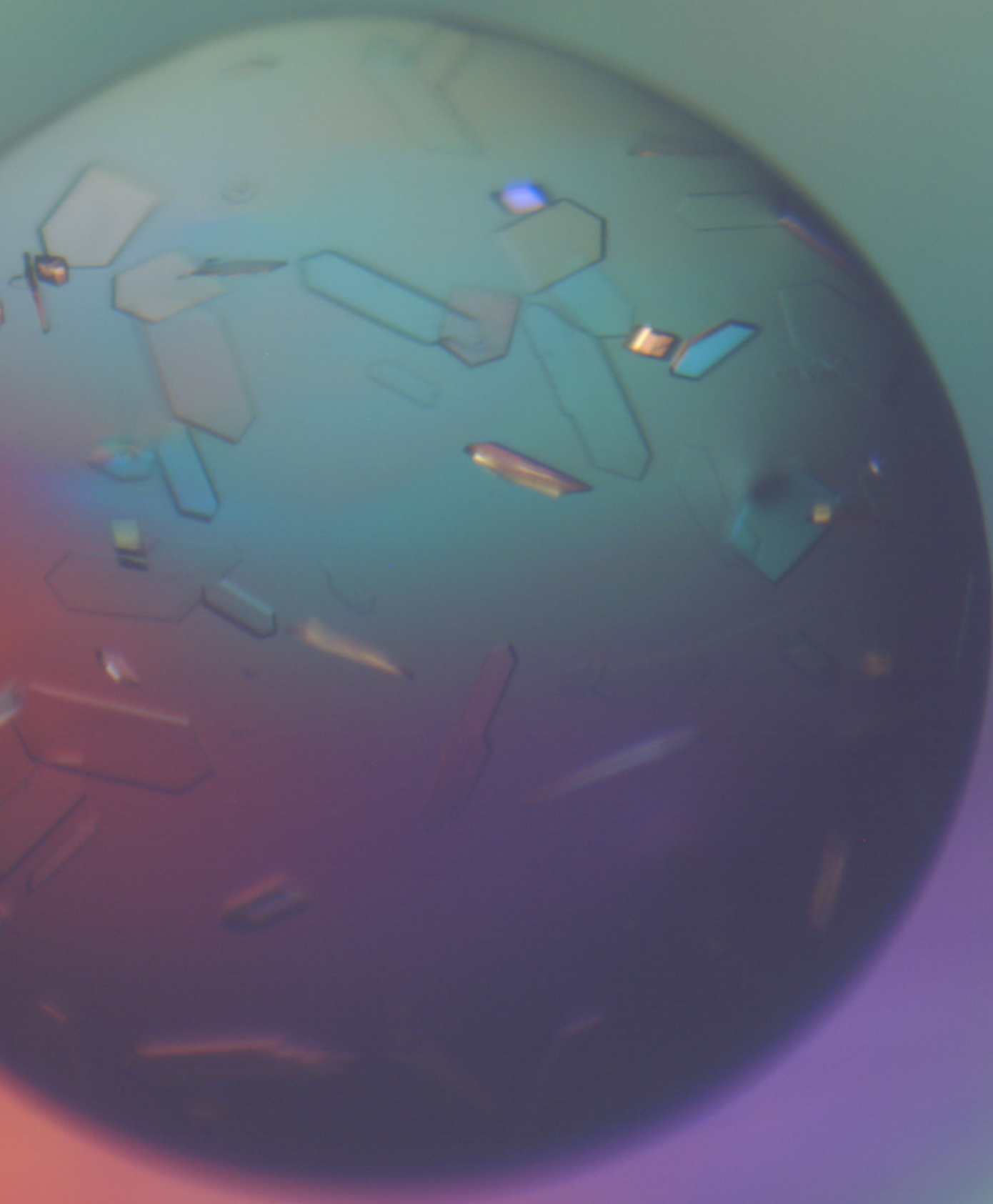
USP7 – ubiquitin-specific-processing protease 7

UTI – urinary tract infection

UV – ultraviolet

Introduction

1



1.1 Antibiotic resistance

The discovery and implementation of the sulfonamide ('sulfa') class of antibiotics (Figure 1.1) in the 1930s heralded the start of the antibiotic revolution, in which many classes of antibiotics were discovered and commercialised, lowering morbidity and mortality from infectious diseases.¹⁻⁵ The phenomenon of antibiotic resistance was, however, observed only shortly after each antibiotic's implementation.⁶ Since this time, antibiotic resistance has eviscerated the pool of effective chemotherapeutic agents for infectious diseases^{7, 8} — regardless of the target or chemical class of the drug — with bacteria evolving with unprecedented ingenuity to counter every antibiotic deployed.^{9, 10} The choice of drug therefore often falls back on those agents which are less effective or more toxic, with severe, life-threatening side effects in some cases.⁷

Antibiotic resistance, originally a natural phenomenon to compete against antibiotic-spreading bacteria, is increasingly populating the micro-biosphere with resistant species as a result of our use and misuse of antibiotics in human, veterinary, and agricultural medicine.^{6, 11} During selective pressure, bacteria harbouring new or pre-existing mutations, or that have incorporated new genetic material, can flourish in the presence of an antimicrobial compound.^{10, 12} Mechanisms of resistance often relate to drug inactivation either metabolically or through efflux.¹⁰ Additionally, mutations can change the drug target, thereby impacting on the binding capacity of a drug, but without significantly compromising binding of the endogenous substrate or bacterial 'fitness', leading to drug resistance.^{10, 13}

Antibiotic resistance has thwarted nearly all successful antibiotic agents and is quickly emerging as one of the most significant health challenges of the century.^{14, 15} The loss of agents from our antibiotic arsenal without appropriate replacement has resulted in warnings of a "post-antibiotic era", in which diseases previously under our control may, once again, go unchecked.^{12, 14} There is a dearth of novel antibiotic agents reaching the market, with only three new classes of antibiotics approved since 1999.¹⁶ The long and difficult task of developing effective antibacterial agents^{9, 17, 18} is sadly not embraced by pharmaceutical companies who prefer the manufacture of drugs for more "lucrative" disease states.^{9, 18, 19} New fast-track government regulations to encourage antibiotic research are thus being implemented.¹⁹⁻²² Due to the inevitable²³⁻²⁵ nature of antibiotic resistance, basic and translational research into new antibiotic agents represents an ongoing area of high priority for current and future health care.

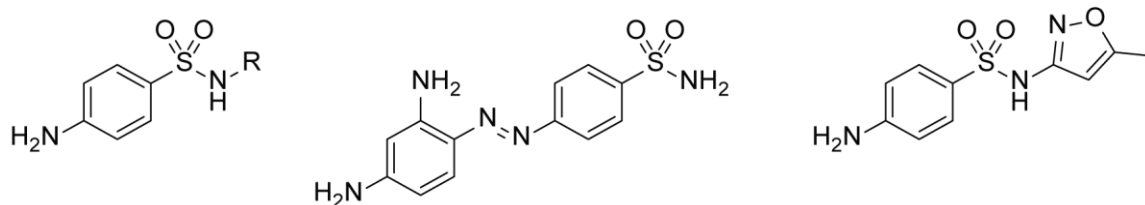


Figure 1.1. Structure of sulfa antibiotics (left), the first developed sulfa drug (the prodrug Prontosil, middle), and the most widely used sulfa drug, sulfamethoxazole (right).

1.2 *Staphylococcus aureus*

The Gram-positive pathogen *Staphylococcus aureus* is a leading cause of bacterial infections worldwide.^{26, 27} Although usually a commensal organism (20–30% of people have a persistent asymptomatic nasal colonisation by *S. aureus*^{28, 29}), invasive infection is responsible for significant morbidity and mortality.³⁰ Initially treated by penicillin, resistance of *S. aureus* to this drug was pandemic in the late 1950s, resulting from the acquisition of β -lactamase genes.^{31, 32} The β -lactamase-resistant drug, methicillin, was used as a treatment, however resistance was already noted in 1961, merely two years after its introduction.³³ Methicillin-resistant *S. aureus* (MRSA) is currently responsible for ~80,000 invasive infections and ~11,000 deaths per year in the US alone.³⁴ While this figure has improved from previous years (an estimated ~111,000 invasive infections occurred in 2005³⁴), MRSA remains a serious healthcare threat; in the United States, *S. aureus* is the number one cause of hospital-acquired (nosocomial) infections.³⁵⁻³⁷ Although endemic in hospitals and healthcare settings worldwide, community-acquired MRSA³⁸ presents its own unique challenges, being able to infect healthy individuals,³⁹ and is present in most developed countries.⁴⁰ The high virulence of community-acquired MRSA, particularly the USA300 strain, results from the production of the cytolytic toxins, α -haemolysin,⁴¹ α -type phenol soluble modulins,⁴² as well as the arginine catabolic mobile element (ACME) that aids bacterial growth and survival in the acidic, low oxygen environment found in pus.⁴³

MRSA has developed resistance to many treatments and few effective chemotherapies remain.⁴⁴⁻⁴⁶ The first-line treatment for serious MRSA infections, vancomycin,⁴⁷ has significant side-effects including nephrotoxicity,⁴⁸ and vancomycin resistance has emerged.⁴⁹⁻⁵² Other treatment options include daptomycin,⁴⁴ linezolid,⁵³ and the combination therapy sulfamethoxazole-trimethoprim,^{54, 55} all of which are met with varying

rates of success in the clinic. In the wake of this and the devastating effects from life threatening MRSA infections, such as bacteraemia or necrotising fasciitis,⁵⁶ the development of novel agents to treat MRSA infections is a health priority.

1.3 *Escherichia coli*

Escherichia coli is a Gram-negative bacterium, and the most studied prokaryote and best characterised model organism.^{57, 58} It has found particular use in the laboratory as a recombinant expression system for proteins due to its low cost, fast growth rate, versatility, high yield, and ease of scale-up⁵⁹⁻⁶¹ — one-third of protein therapeutics are produced in *E. coli*.⁵⁹ *E. coli* is, however, incapable of eukaryotic post-translational modifications and some, generally larger, proteins will simply fail to express or will express as insoluble inclusion bodies.⁶¹⁻⁶³ *E. coli* is able to thrive using glucose and ammonium as a sole carbon and nitrogen source, respectively, making it the first organism of choice as a suitable expression system to ¹⁵N/¹³C-isotope-label a protein, for use in nuclear magnetic resonance (NMR) spectroscopy protein studies.⁶⁴

E. coli usually exists as a harmless component of human gut flora, rarely causing disease.⁵⁷ Although 100-fold to 1000-fold less numerous than obligate anaerobes, *E. coli* is the predominant aerobe in the gastrointestinal tract.⁶⁵ However, many pathotypes of *E. coli* do exist, generally causing either enteric/diarrheal disease, urinary tract infections (UTIs), or sepsis/meningitis,⁶⁶ and outbreaks are common in both developed and developing countries.⁶⁷ Diarrheagenic *E. coli* infection is less common in developed countries but remains a serious issue in children in developing countries.⁶⁷⁻⁶⁹ Diarrheal disease is responsible for 10% of worldwide deaths of children under the age of 5,⁷⁰ of which *E. coli* is a major contributor.^{67, 69} Treatment includes the agents rifamixin and ciprofloxacin.^{71, 72} Significantly, in May 2016 a 49 year old woman from Pennsylvania presented with an *E. coli* urinary tract infection harbouring the mcr-1 gene that confers resistance to colistin, usually the last line of defence against multidrug-resistant Gram-negative bacteria.⁷³

Intestinal *E. coli* acts to protect the host against pathogens through colonisation resistance,⁷⁴⁻⁷⁶ and disruption of intestinal flora by an antibiotic regimen has been demonstrated to promote both secondary infections⁷⁷⁻⁷⁹ and antibiotic resistance.⁸⁰ The

development of species-selective antibacterial agents (in this case those lacking potency against *E. coli*) is therefore being increasingly advocated.^{24, 81}

1.4 The folate pathway

Folic acid or folate (vitamin B9) is a crucial cofactor in the synthesis of various amino acids, purines, thymidine, and panthenoic acid, and is essential for DNA replication and life.^{82, 83} Humans and higher eukaryotes depend on both passive and active uptake of dietary folate, however, other organisms, including bacteria and plants, synthesise folate *de novo*.^{83, 84} The absence of various folate pathway enzymes in humans makes them attractive drug targets. Two of the enzymes in this pathway, dihydropteroate synthase (DHPS) and dihydrofolate reductase (DHFR), are known drug targets, DHPS being the target of the aforementioned sulfonamides.^{85, 86} DHFR is the target of trimethoprim and the prospective antibiotic iclaprim.⁸⁷ The folate biosynthetic pathway is summarised in Figure 1.2 and involves the sequential processing of guanine triphosphate to tetrahydrofolate (THF) via the enzymes GTP cyclohydrolase (GTPCH), nudix phosphatase, dihydroneopterin aldolase (DHNA), 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), DHPS, dihydrofolate synthase (DHFS), and DHFR.⁸⁶ The structural overlap of the folate pathway substrates, and known antifolates, has been noted,^{85, 86, 88} and an inhibitor has even been observed in both enzymes of a crystal structure of the bifunctional HPPK-DHPS complex from *Francisella tularensis*,⁸⁹ indicating the potential to develop a multi-enzyme-targeting inhibitor.

In *S. aureus*, 351 genes were identified to be important for survival and growth, of which six genes of the folate pathway were identified to be essential.⁹⁰ Of these, the most relevant to this thesis are *folK* and *folP*, coding for the enzymes HPPK and DHPS, respectively, implicating these enzymes to be attractive antibacterial targets.

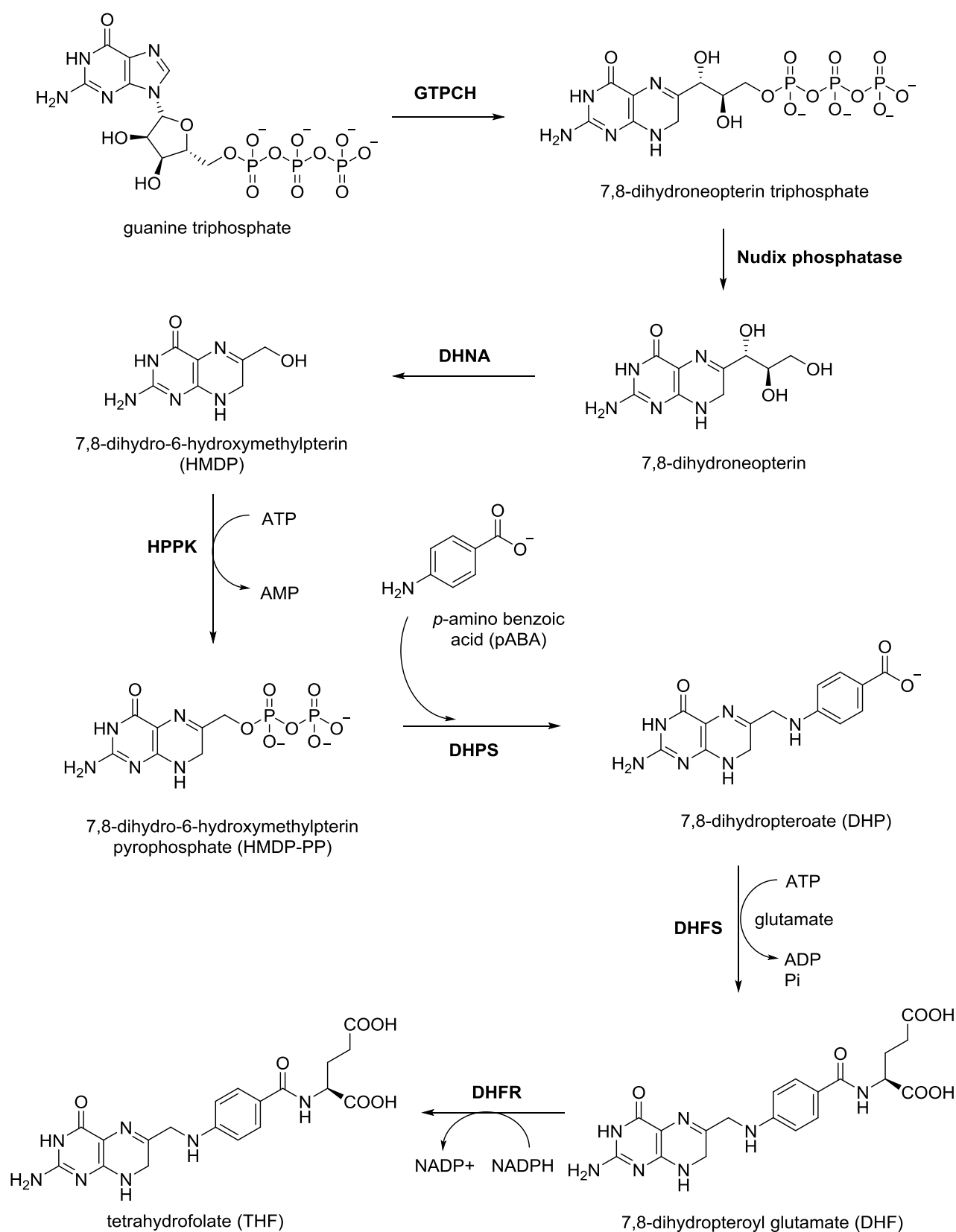


Figure 1.2. The folate biosynthetic pathway. The sequential enzymes HPPK and DHPS are the focus of this thesis.

1.5 HPPK

HPPK catalyses the transfer of a pyrophosphate molecule from adenosine triphosphate (ATP) to the substrate 7,8-dihydro-6-hydroxymethylpterin (HMDP), forming 7,8-dihydro-6-hydroxymethylpterin pyrophosphate (HMDP-PP).⁹¹ HPPK is small (18 kDa), possessing an α - β - α thioredoxin-like fold.^{92, 93} Crystal structures of HPPK have been solved for a variety of species,⁹³⁻⁹⁷ including *S. aureus* (SaHPPK).⁹⁸ HPPK is typically a monofunctional monomeric enzyme but does exist as part of a multi-functional enzyme in some species, either as: bifunctional DHNA-HPPK (e.g. *Streptococcus pneumoniae*⁹⁵), bifunctional HPPK-DHPS (e.g. *Francisella tularensis*^{89, 97}), or trifunctional DHNA-HPPK-DHPS (e.g. *Saccharomyces cerevisiae*⁹⁶).

The core of the protein remains rigid throughout catalysis but there are three loop regions (L1–3, residues 12–14, 47–51, and 82–94, respectively, SaHPPK numbering), in particular L3, which undergo significant conformational changes (Figure 1.3).^{91, 99} In the apo-enzyme, these loops are markedly more mobile than in the bound-enzyme.⁹⁹ Catalysis occurs via an ordered mechanism in which the ATP cofactor binds first — this causes L3 to extend away from the catalytic site, L2 remaining mobile.^{91, 100} The pterin substrate then binds and L3 closes over the active site; the resulting hydrogen bond network between the three loops rigidifies the enzyme, forming the Michaelis complex, and allowing for catalysis to occur (Figure 1.3C).^{91, 100} Subsequent to pyrophosphate transfer, adenosine monophosphate (AMP) dissociates followed by dissociation of the pterin-pyrophosphate product molecule, the rate-limiting step.^{101, 102} HPPK also has ATPase activity and will cleave the γ -phosphate of ATP, as was demonstrated by an adenosine diphosphate (ADP) molecule being observed in a crystal structure of HPPK soaked with ATP.⁹⁹ The non-reactive α,β -methyleneadenosine 5'-triphosphate (AMPCPP) has been proposed as the most appropriate analogue of ATP,¹⁰³ and has been used in HPPK structures from three species.^{89, 104, 105}

Much kinetic and binding data has been acquired on HPPK from *E. coli* (EcHPPK).^{38, 101, 102, 106-108} The binding properties of a range of substrate, product, and analogue compounds towards EcHPPK are summarised in Table 1.1. The kinetic constants for EcHPPK substrates are summarised in Table 1.2. Data on the binding of substrate and substrate analogue molecules to SaHPPK was published by Chhabra et al.⁹⁸ and is displayed in Table 1.3. Of note is the significantly reduced binding of the HMDP substrate in the absence of

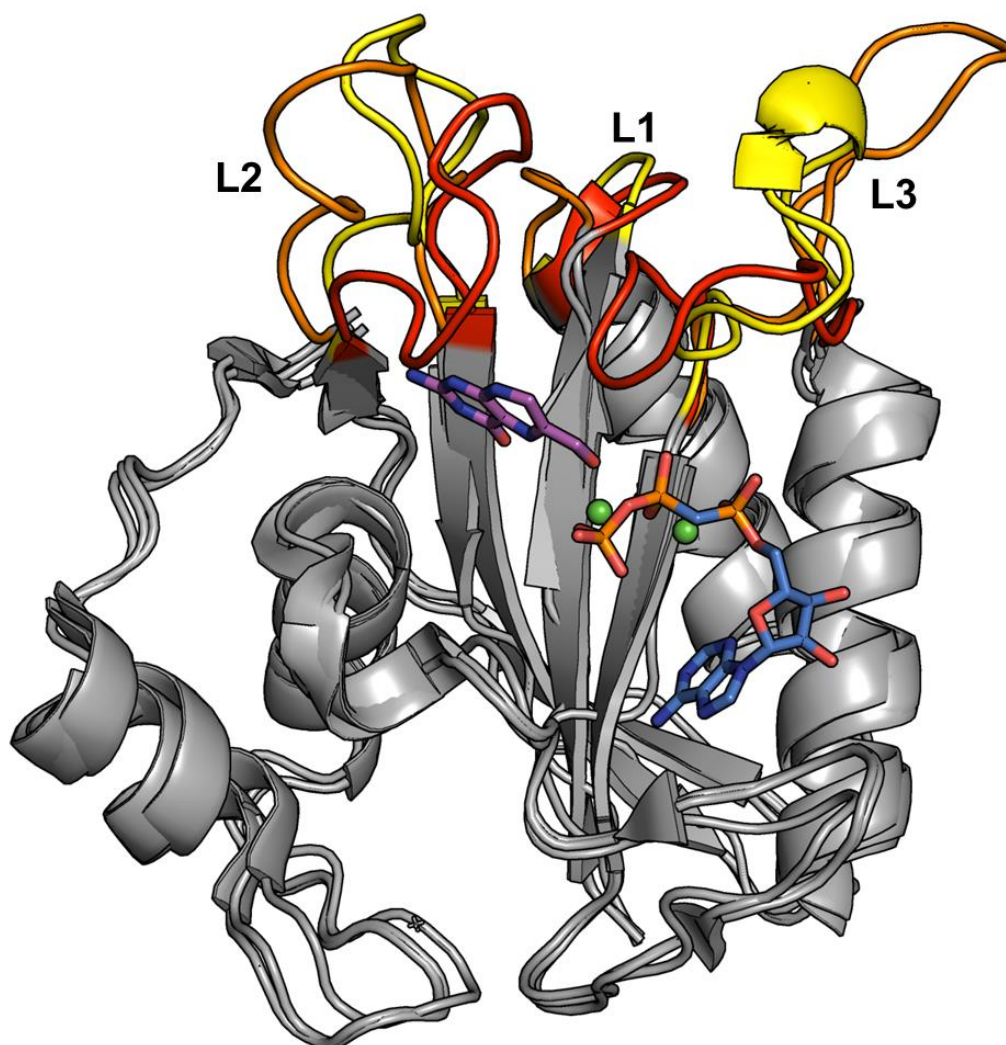
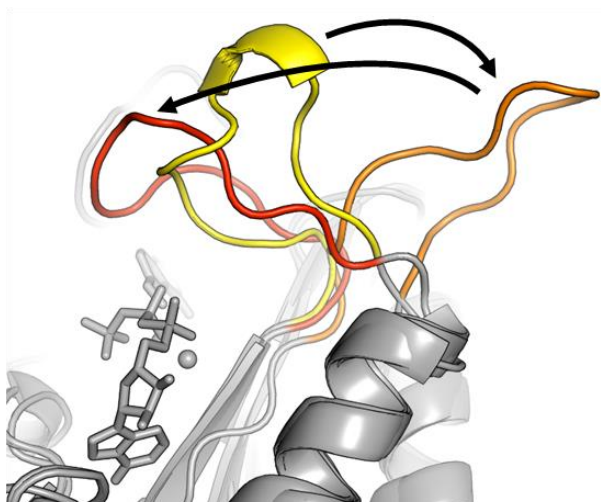
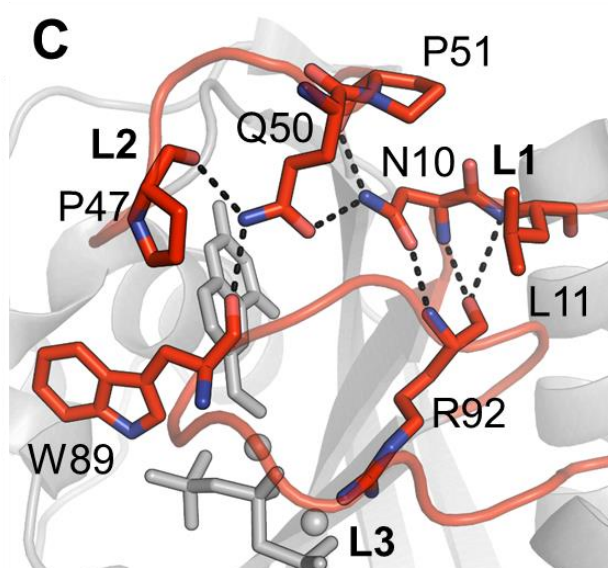
A**B****C**

Figure 1.3. (A) Overlay of (*E. coli*) HPPK structures obtained for various stages of the catalytic cycle with L1–3 highlighted. The apo- (yellow, PDB ID: 1HKA⁹³), cofactor-bound (orange, PDB ID: 1EQ0⁹⁹), and cofactor and substrate-bound (red, PDB ID: 1Q0N¹⁰⁴) enzyme structures are shown. The pterin substrate is displayed in purple, AMPCPP is displayed in blue, and the catalytic magnesium ions are displayed in green. (B) The movements of L3 during catalysis. (C) The hydrogen bond network between residues of L1–3 in the HPPK ternary complex.

cofactor,^{98, 102} likely reflecting the fact that substrate/ligand binding in both pterin- and cofactor-binding pockets is needed to stabilise and rigidify the enzyme.¹⁰⁰ The ATP substrate binds tighter than either ADP or AMP (~15-fold and 40-fold, respectively for EcHPPK, and ~17-fold and ~110-fold, respectively for SaHPPK)^{98, 102, 106, 107} and ATP binding is itself notably weaker in the absence of Mg²⁺ (~11-fold for EcHPPK).^{106, 107} Changing from ATP to GTP reduces binding ~140-fold in EcHPPK.^{102, 106, 107} The substrate analogue AMPCPP is noted to bind markedly tighter than ATP (7-fold in SaHPPK and 8–46-fold in EcHPPK).^{98, 102, 106} Kinetic data obtained has suggested that ATP binding occurs first in a slower step, followed by the rapid addition of HMDP, with product release being rate limiting.^{101, 102, 106}

Despite the range of structural and kinetic data on HPPK, few inhibitors have been developed, and there are currently no drugs that are known to target HPPK. Compounds **1** and **2** (Figure 1.4) were designed as analogues of the pterin substrate. These were the first HPPK inhibitors to be developed, and have been crystallised in EcHPPK, despite affinity data not being reported.^{94, 109, 110} A series of bi-substrate inhibitors have also been developed and tested on EcHPPK.^{111–114} Initially these consisted of a pterin moiety connected to an adenosine via a linker consisting of 2–4 phosphate groups (HP_nA, Figure 1.4). *K_d* values were 4.25 μM, and 0.47 μM for linker lengths of 3 and 4 phosphate groups, respectively, and IC₅₀ values were >100 μM, 1.27 μM, and 0.44 μM, for linker lengths of 2–4 phosphate groups, respectively. The longest linker was thus determined to be optimal. Subsequent work focused on attempts to improve the physicochemical properties of the linker through replacement of the phosphate groups, an example being compound HP-19 (Figure 1.4). However, a loss of binding and inhibition (≥ 5-fold) was associated with these modifications.

Table 1.1. Binding affinities of various substrate, products, and analogue compounds for EcHPPK.

Ligand Name	K_d (μM)	ΔG ($\text{kcal}\cdot\text{mol}^{-1}$)	Reference
ATP (with Mg^{2+})	4.5 ± 0.6	-7.3	106
ATP (with Mg^{2+})	2.6 ± 0.06	-7.6	102
ATP (without Mg^{2+})	38 ± 2.7	-6.0	107
ADP (with Mg^{2+})	55 ± 3.3	-5.8	107
AMP (with Mg^{2+})	48 ± 5	-5.9	106
AMP (with Mg^{2+})	140 ± 11	-5.3	107
Adenosine (with Mg^{2+})	210 ± 14	-5.0	107
GTP (with Mg^{2+})	340 ± 50	-4.7	106
GTP (with Mg^{2+})	680 ± 30	-4.3	107
AMPCPP (with Mg^{2+})	0.077 ± 0.006	-9.7	102
AMPCPP (with Mg^{2+})	0.45 ± 0.04	-8.6	106
HMDP (with Mg^{2+} and AMPCPP)	0.036	-10.1	106
HMDP (with Mg^{2+} and AMPCPP)	0.17 ± 0.01	-9.2	102
HMDP (without Mg^{2+})	110 ± 17	-5.5	102
HMDP-PP (with Mg^{2+})	0.2	-9.1	101

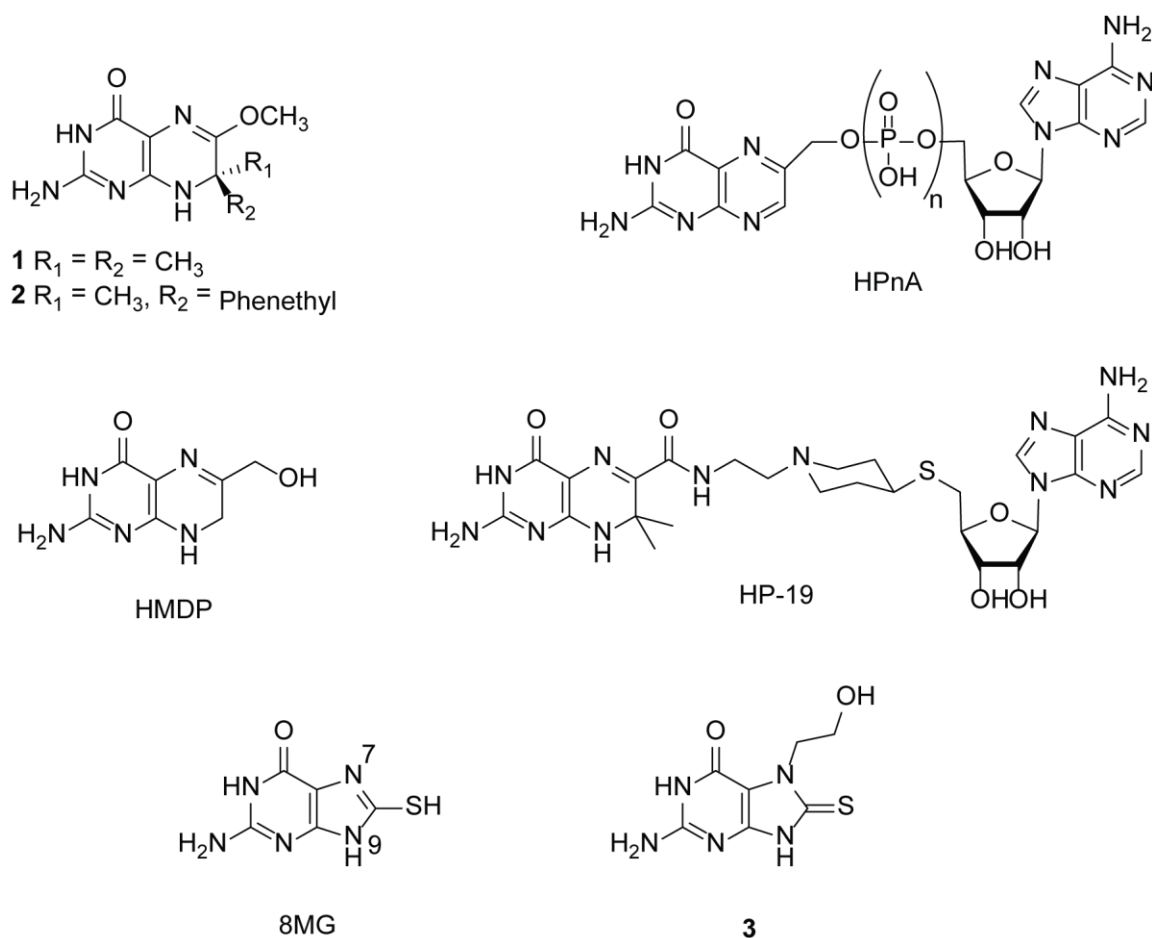
Table 1.2. Kinetic constants (K_m) for EcHPPK substrates.

Substrate Name	Method	K_m (μM)	Reference
HMDP	Steady state kinetics	0.60 ± 0.01	106
HMDP	Radioactivity-based assay	1.6 ± 0.4	108
ATP	Steady state kinetics	3.4 ± 0.02	106
ATP	Radioactivity-based assay	17 ± 3	108

Table 1.3. Binding affinities of substrate and substrate analogues to SaHPPK.

Ligand ^{ab}	K_d (μ M) ITC	K_d (μ M) SPR
ATP	31.0 ± 4.5	45 ± 2
ADP	–	760 ± 16
AMP	–	4900 ± 1100
AMPCPP	3.1 ± 1.2	7.7 ± 0.4
HMDP	–	100 ± 12
HMDP (with AMPCPP ^c)	4.0 ± 1.2	3.6 ± 0.3

^aChhabra et al.⁹⁸ ^bWith 10 mM MgCl₂. ^c1 mM AMPCPP.

**Figure 1.4.** Structures of the HPPK substrate (HMDP) and selected HPPK inhibitors. Positions N⁷ and N⁹ of 8MG are identified.

Work by Chhabra et al.⁹⁸ on HPPK inhibitors involved screening a small fragment library of pterin analogues, against SaHPPK, identifying 8-mercaptoguanine (8MG, Figure 1.4) as a compound that bound ($K_D = 12.8 \mu\text{M}$) and inhibited ($IC_{50} = 41 \mu\text{M}$) the enzyme. Chemical extension from the N⁷ and N⁹ positions of 8MG did not prove beneficial, the N⁷-functionalised compound **3** (Figure 1.4) possessing an affinity of $12.3 \mu\text{M}$, and the S⁸-position was deemed the most appropriate site for extension in further studies.¹¹⁵ Crystal structures of 8MG and **3** complexed to SaHPPK were determined.^{98, 115} In these, the compounds bound in a similar fashion to the pterin substrate, possessing a hydrogen bond network to residues 43–46 and N56, with π -stacking between residues F54 and F123 (Figure 1.5). Prior to this thesis, cofactor-bound and ternary complexes of SaHPPK had not been crystallised and represented milestone gaps for structure-based drug design of inhibitors of SaHPPK.

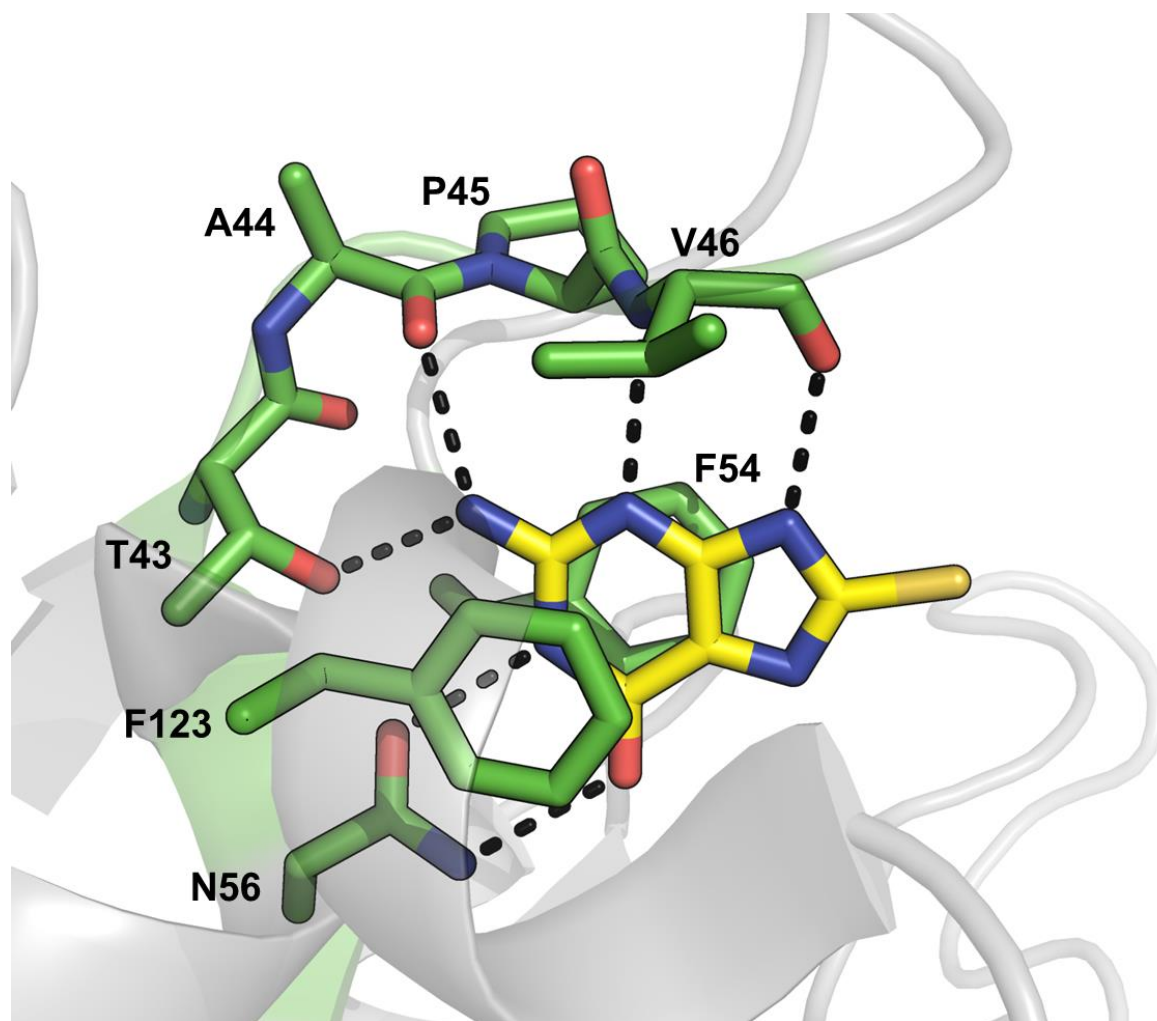


Figure 1.5. Binding of the hit fragment 8MG to SaHPPK (PDB ID: 3QBC⁹⁸). Key residues are highlighted in green, 8MG is displayed in yellow, and hydrogen bonds are displayed as dashed black lines.

1.6 DHPS

DHPS is the HPPK-adjacent, downstream enzyme in the folate pathway that catalyses the condensation of HMDP-PP and *para*-aminobenzoic acid (pABA) to form 7,8-dihydropteroate (DHP).¹¹⁶ In bacteria DHPS typically exists as a ~60 kDa homodimer,¹¹⁶⁻¹¹⁹ although, as discussed above, higher-order association with the preceding enzymes does occur (i.e. HPPK-DHPS⁹⁷ and DHNA-HPPK-DHPS⁹⁶). Each DHPS monomer comprises a single domain with a “TIM-barrel” (α/β)₈-folding topology, and houses a deep pterin-site binding-pocket and adjacent diphosphate- and pABA-binding regions (Figure 1.6).¹¹⁶⁻¹¹⁸ Catalysis occurs via an S_N1 mechanism.¹¹⁹ Upon binding of the substrate (HMDP-PP), the pyrophosphate is cleaved. During this step, the HMDP carbocation intermediate is stabilised via delocalisation of the charge in the adjacent double bond of the pterin ring, which is facilitated by interactions with the conserved binding-site residues D96 and D185 (*E. coli* DHPS (EcDHPS) numbering). Binding of pABA and subsequent condensation with this intermediate leads to the dihydropteroate product.

The sulfonamide drugs are substrate analogues that target the pABA-binding pocket of DHPS.¹¹⁸ However, mutations in the loop residues that form the pABA pocket have often led to sulfonamide resistance.¹³ The flexibility of this region has been linked to the development of resistance.¹²⁰ In comparison, the pterin binding pocket is highly rigid, in addition to being highly conserved between species.¹¹⁸ Targeting this binding pocket has therefore been suggested as an approach to achieve broad spectrum activity and a reduced susceptibility to resistance-bearing mutations.^{118, 121, 122}

Work on pterin-site inhibitors of DHPS began in the 1970s with a series of 4,5-dioxo-1,4,5,6-tetrahydropyrimido[4,5-*c*]pyridazines developed by Burroughs Wellcome Co. (London, UK).^{123, 124} A decade later various isocytosine derivatives were synthesised and tested against EcDHPS (Figure 1.7).¹²⁵ Of these, compound **4** displayed the best inhibition — 1.6 μ M in a radiolabelled-pABA assay. Using **4** as a scaffold, with extension from the methylamine, a series of different length hydrocarbon linkers attached to a phenyl ring were tested. Although tolerated, no improvement in inhibition was achieved for these groups, the best inhibition being obtained for a four-carbon linker (compound **5**, Figure 1.7, IC₅₀ = 2.5 μ M) or seven-carbon linker (IC₅₀ = 1.4 μ M). A following study looked at analogues of **5**, with a sulfonamide-toluene derivative (compound **6**) displaying the best inhibitory activity (0.7

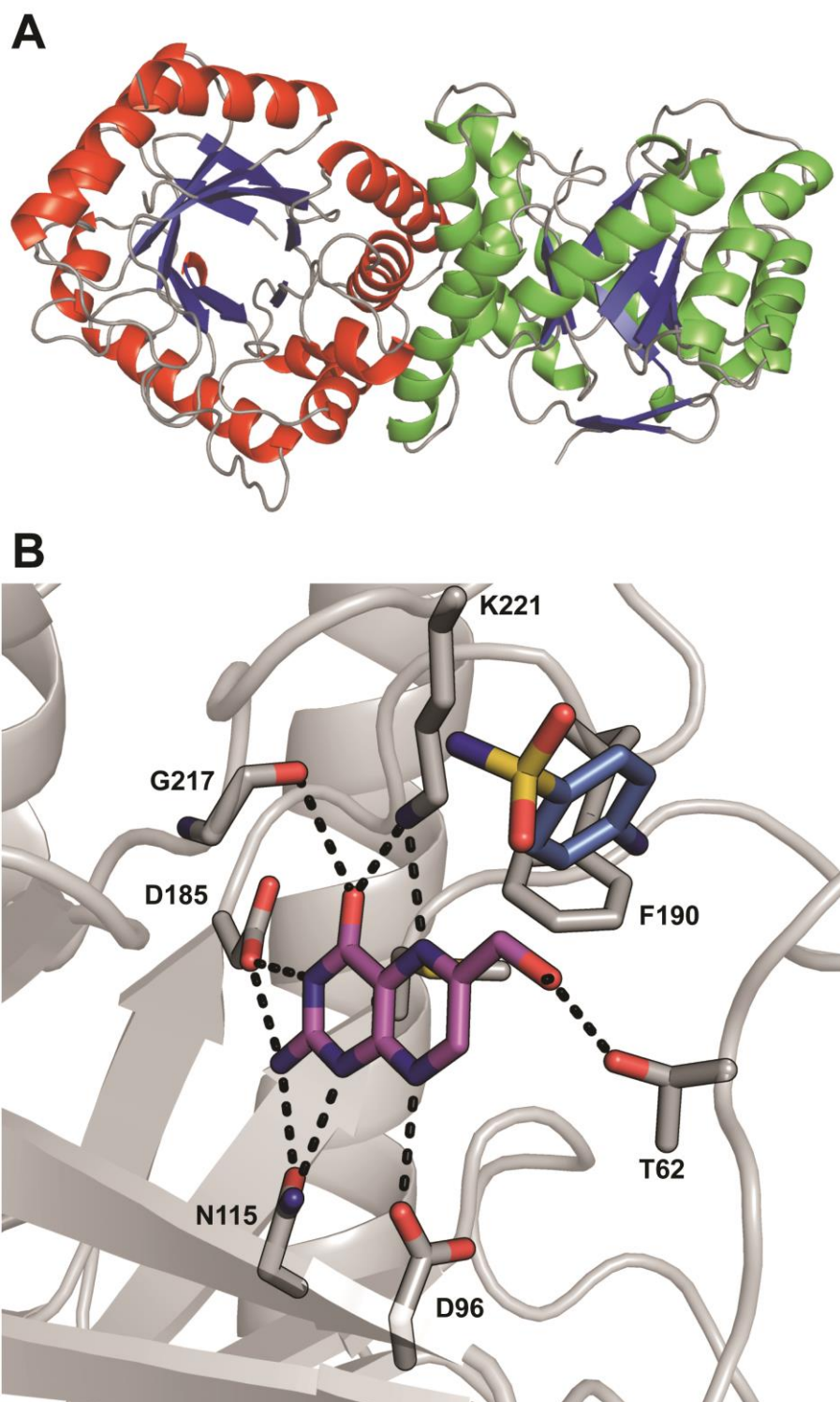


Figure 1.6. (A) Structure of the (*E. coli*) DHPS dimer. Helices, sheets, and loops are coloured red/green, blue, and grey, respectively. (B) The pterin-binding site of DHPS (PDB ID: 1AJ0¹¹⁸). The pterin-intermediate state analogue is coloured purple and the sulfonamide (i.e. sulfanilamide) is coloured blue. Hydrogen bonds from the protein to the pterin moiety are displayed as dashed black lines.

μM).¹²⁶ One of the compounds from the initial study¹²⁵ was later crystallised in DHPS from *Bacillus anthracis* (BaDHPS), confirming that the pterin site was being targeted.¹¹⁷

A study using virtual screening identified and characterised the binding of several novel ligands to BaDHPS.¹²¹ Two compounds from this study, **7** and **9** (Figure 1.7), gave IC_{50} values of $19.8 \mu\text{M}$ and $87.1 \mu\text{M}$ respectively. 8MG was also shown to bind to BaDHPS, as evidenced by a crystal structure, although no inhibition of the enzyme was observed. Despite various novel compounds and scaffolds being found, no compound possessed better inhibitory efficacy than the previously reported **4** ($\text{IC}_{50} = 8.0 \mu\text{M}$).¹²¹ A study developing compound **7** was subsequently performed.¹²² Of the various derivatives synthesised and tested, a simple *N*-demethylated analogue (**8**) proved the most effective, with an IC_{50} of $11 \mu\text{M}$ and K_D of 76 or 120 nM, as determined by isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR), respectively.¹²² Compound **9** is similar to the catalysis product, DHP. A recent study developed a series of analogous compounds (**10**) featuring a two-atom linker and a sulfonamide group in place of the carboxylic acid moiety, to resemble a pterin-sulfonamide conjugate.¹²⁷ A fairly “flat” structure-activity relationship (SAR) was observed for the compounds tested, with a less than 2-fold difference for seven of the eight compounds tested ($\text{IC}_{50} \sim 20 \mu\text{M}$ in a fluorescence polarisation assay).

Additionally, several compounds (e.g. **11**) intended to simultaneously target the pterin, phosphate, and pABA binding-sites were developed.¹²⁸ These compounds displayed a large loss of potency compared to those discussed above. This may reflect the fact that these were designed to mimic an $\text{S}_{\text{N}}2$ -mechanism transition state, in which pyrophosphate elimination is concerted with pABA addition,¹²⁸ prior to it being established that DHPS catalysis actually proceeds through an $\text{S}_{\text{N}}1$ mechanism, in which pyrophosphate elimination occurs prior to pABA addition (pABA reacting with a charged HMDP intermediate).¹¹⁹ Finally, inhibitors targeting an allosteric site of DHPS have been discovered, the fragment **12** inhibiting DHPS from the three species tested (BaDHPS and DHPS from *Staphylococcus aureus* and *Yersinia Pestis*).¹²⁹ Crystal structures from this study revealed binding at the dimer interface for several compounds, and a K_d of $187 \mu\text{M}$ for BaDHPS binding was determined by SPR.¹²⁹

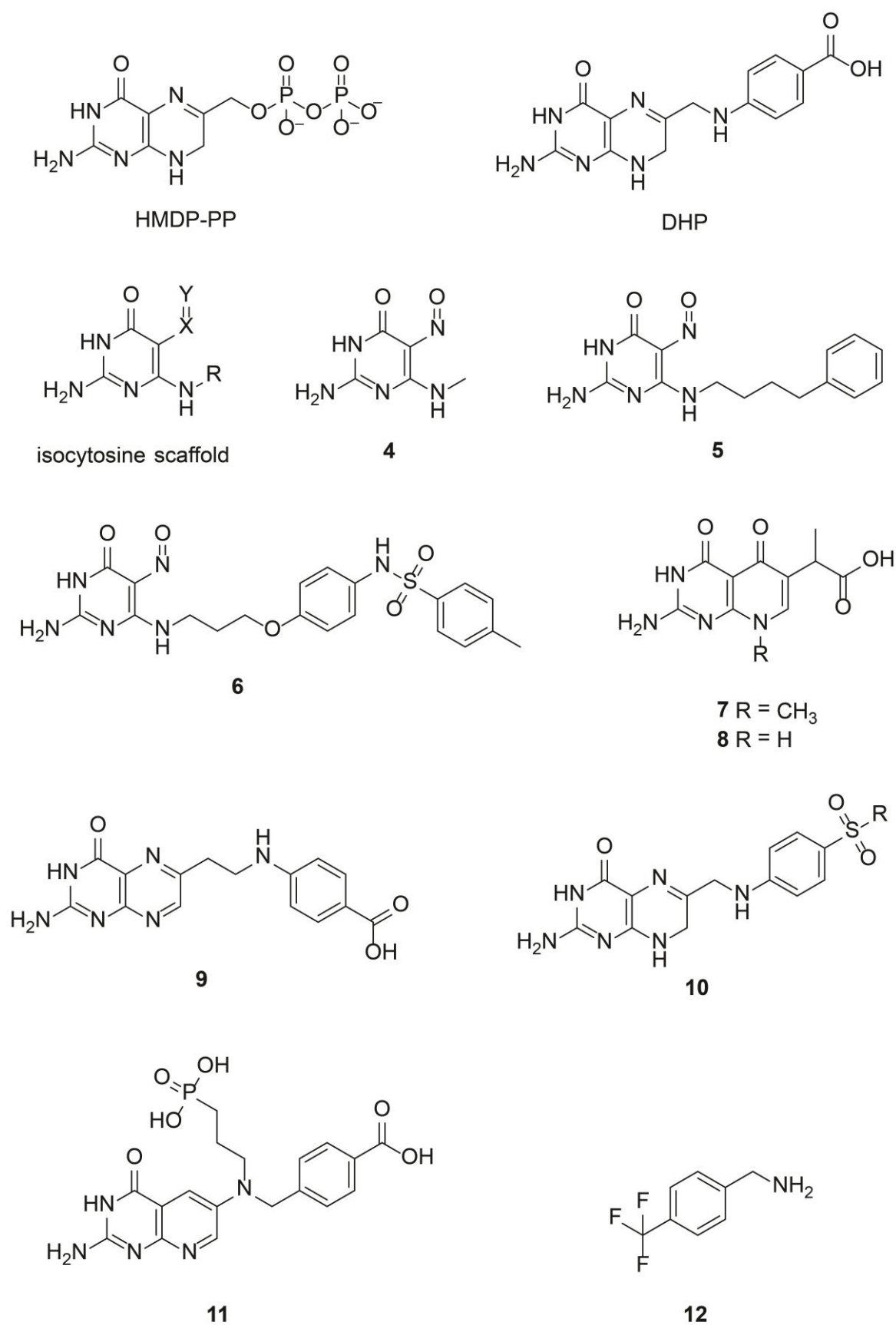


Figure 1.7. Structures of the DHPS substrate (HMDP-PP), product (DHP), and selected DHPS inhibitors.

1.7 Hit development in drug discovery

The advent of high-throughput screening (HTS) in the 90s, in which large libraries of compounds were tested against prospective targets, did not provide the returns expected. A high (and expensive) rate of attrition, attributable in part to an overemphasis on potency, afflicted all major pharmaceutical companies.¹³⁰ Compounds discovered from these methods typically lacked synthetic amenability and had poor pharmacokinetics outcomes. Lipinski's seminal paper highlighted the need to address the poor physicochemical properties of the compounds presented in chemical libraries.¹³¹ The "rule of 5"¹³¹ proposed that poor absorption or permeation would occur when compounds possessed: a molecular weight > 500, logP > 5, hydrogen-bond donors > 5, and hydrogen-bond acceptors > 10. Other key rules that extended this concept of "druglikeness" included an aim towards rotatable bonds ≤ 10 , and polar surface area $\leq 140 \text{ \AA}^2$.¹³² Increasing consideration of these concepts in drug development improved the success rate — clinical (Phase I–III trials) attrition due to pharmacokinetics/bioavailability reduced from 40% to 8% between 1991 and 2000.¹³³ The related concept of "lead-likeness",¹³⁴ in which compounds are expected to increase in lipophilicity and molecular weight during development, has been incorporated into screening libraries and improved success rates.¹³⁵ Consideration of other concepts, including diversity and synthetic amenability, has improved the quality of HTS libraries and it today stands as a proven method for drug discovery.¹³⁵

1.7.1 Fragment-based drug discovery

In 2011, 15 years after the seminal "SAR by NMR"¹³⁶ paper presented the fragment-based drug discovery (FBDD) concept, the first fragment-derived drug, Zelboraf, was approved by the FDA as a cancer therapeutic.^{137, 138} FBDD considers the idea of lead-likeness, but extends it further to a "rule of 3", whereby compounds are promoted that possess a molecular weight < 300, hydrogen-bond donors ≤ 3 , hydrogen-bond acceptors ≤ 3 , and cLogP ≤ 3 .^{139, 140} While useful fragments are found outside these parameters, the crux of this guideline is a shift towards simpler molecules that can better sample chemical space.^{140, 141} The compounds that fill fragment-based screening libraries are expected to bind or inhibit their targets with only low affinity (typically high μM to low mM).¹⁴² The focus, rather than on affinity itself, is instead on "ligand efficiency",¹⁴³ although few, the protein-ligand interactions are more optimised than those typically returned from an HTS

campaign.¹⁴² Additionally, the hit rate is typically higher than that of HTS due to a reduced chemical complexity.^{144, 145}

FBDD derives from the understanding that ligands tend to increase in size and lipophilicity as they are developed; smaller, more efficient ligands that possess suitable sites for synthetic elaboration are thus prioritised.^{142, 146} Typical methods used in FBDD include fluorescence-based thermal shift,¹⁴⁷ NMR spectroscopy,¹⁴⁸ X-ray crystallography,¹⁴⁹ SPR,¹⁵⁰ and ITC.¹⁵¹ Once hit fragments are found, preferably with a determined binding mode, three key approaches can be employed to improve potency: merging, linking, or growing.¹⁴⁴ Fragment merging and linking occur when fragments that display overlapping and non-overlapping binding modes, respectively, are combined. Linking can produce impressive results¹⁵² but is less reliable and commonplace than fragment growing, in which a single binding site is discovered and fragments are synthetically elaborated to probe additional interactions, and improve potency.¹⁴² Such leads derived from FBDD are typically observed to have improved physicochemical properties compared to those from HTS methods, as exemplified in Figure 1.8.¹⁵³

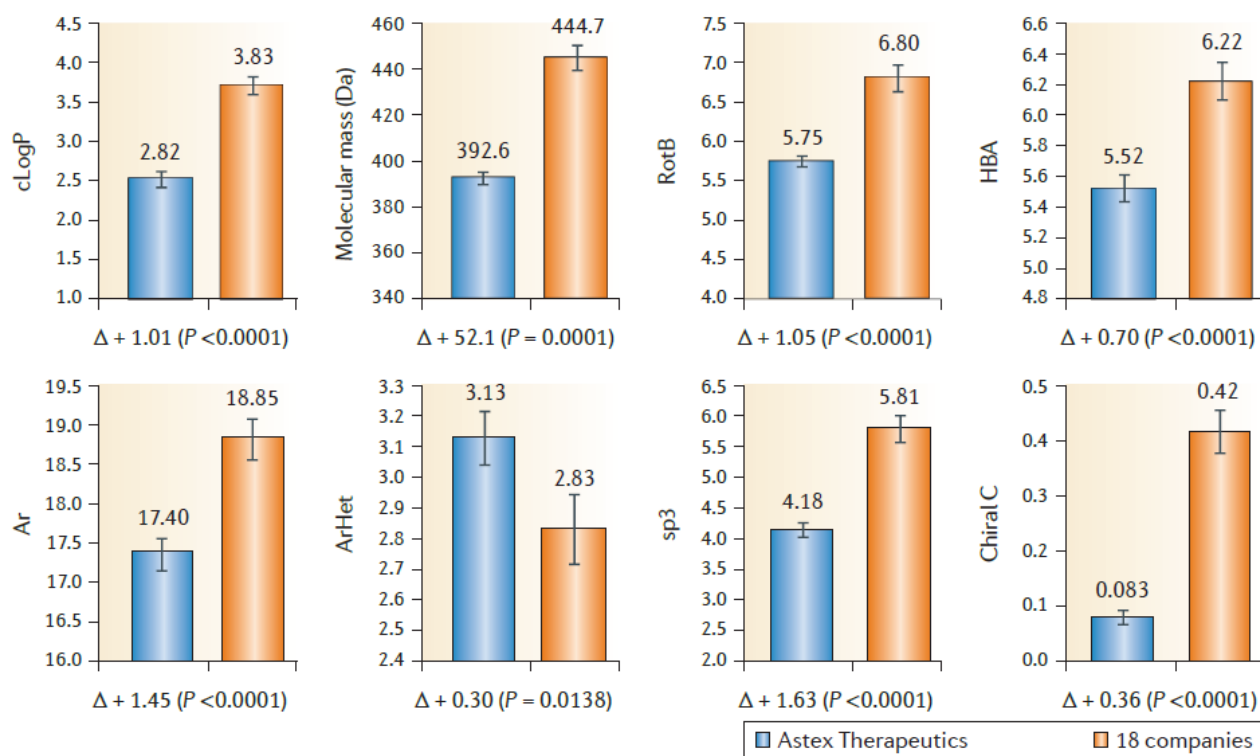


Figure 1.8 The impact of fragment-based drug discovery on physicochemical properties. Significantly more “druglike” molecules are observed from Astex Therapeutics, a company that focuses only on fragment-based drug discovery using X-ray screening methods. The

figure shows matched-pair analyses of molecular properties for the shared single-gene targets from Astex Therapeutics and the 18 companies with ≥ 50 patented compounds. Error bars are \pm standard errors of the mean values. Ar, aromatic atom carbon count; ArHet, aromatic atom non-carbon count; Chiral C, chiral carbon count; HBA, hydrogen-bond acceptor count; sp³, tetrahedral carbon atom count; RotB, rotational bond count. Adapted with permission from Leeson et al.¹⁵⁴ Nature Publishing Group.

1.7.1.1 Structure-based drug design

Knowing the interactions that a target makes with ligand molecules can be useful during compound elaboration and optimisation. As might be expected, and illustrated in Figure 1.9, the ability to acquire NMR or X-ray crystal structures on prospective compounds greatly improves the likelihood of successfully developing a hit to a lead.¹⁵⁵ While other biophysical techniques have their place in structural biology, notably cryo-electron microscopy,¹⁵⁶ NMR spectroscopic and X-ray crystallographic methods predominate, and are the key methods used in this thesis.

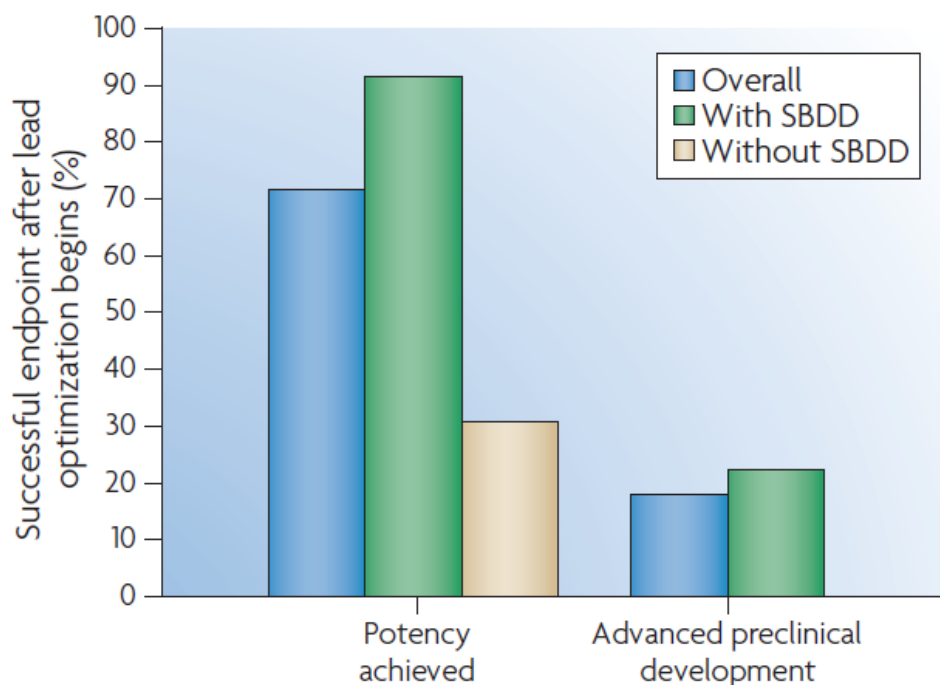


Figure 1.9. The importance of structural information to fragment-based drug design. The success in achieving potent ($IC_{50} < 100$ nM) inhibitors or compounds suitable for advanced preclinical development after a lead optimisation programme was initiated is compared with

and without structure-based drug design (SBDD). The analysis includes synthetic optimisation programmes on 20 leads from fragment-based screening (14 of which received X-ray- or NMR-based structural support). Adapted with permission from Hajduk et al.,¹⁵⁵ Nature Publishing Group.

1.8 Scope of the thesis

Many enzymes of the folate biosynthesis pathway represent attractive antibacterial targets. HPPK and DHPS are two such enzymes and have thus been selected for study in this thesis. The pterin substrate-binding pockets of these enzymes are not known targets of any drugs. Both pockets are highly rigid and conserved, characteristics that are predicted to attenuate antibiotic resistance occurring.^{104, 118} Structural data of HPPK and DHPS have been obtained, in complex with substrates and ligands, providing a foundation for rational structure-guided drug design of hit compounds. This thesis focuses on the fragment 8MG, a compound that has been shown to bind both enzymes, and its elaboration into more potent lead compounds. The work presented herein aims to screen ligands against SaHPPK, EcHPPK, and EcDHPS and to investigate the binding of these ligands using NMR and X-ray crystallography. The structure-activity relationships obtained, and the insights provided by the structural techniques, facilitated the development of these ligands, with an eye on improving potency and drug-likeness. Figure 1.10 shows an outline of the general process used throughout this project. Briefly, proteins are optimised in regards to their expression and stability. Compounds synthesised by collaborators are tested for binding against these proteins (typically via SPR, but also ITC and differential scanning fluorimetry (DSF)). Both experimental (NMR and X-ray crystallography) and *in silico* (molecular docking) determinations of binding mode are performed to understand the mechanisms behind the affinities observed. Insights thus gained are used to direct future synthetic efforts. A detailed description of each technique is provided in Chapter 5.

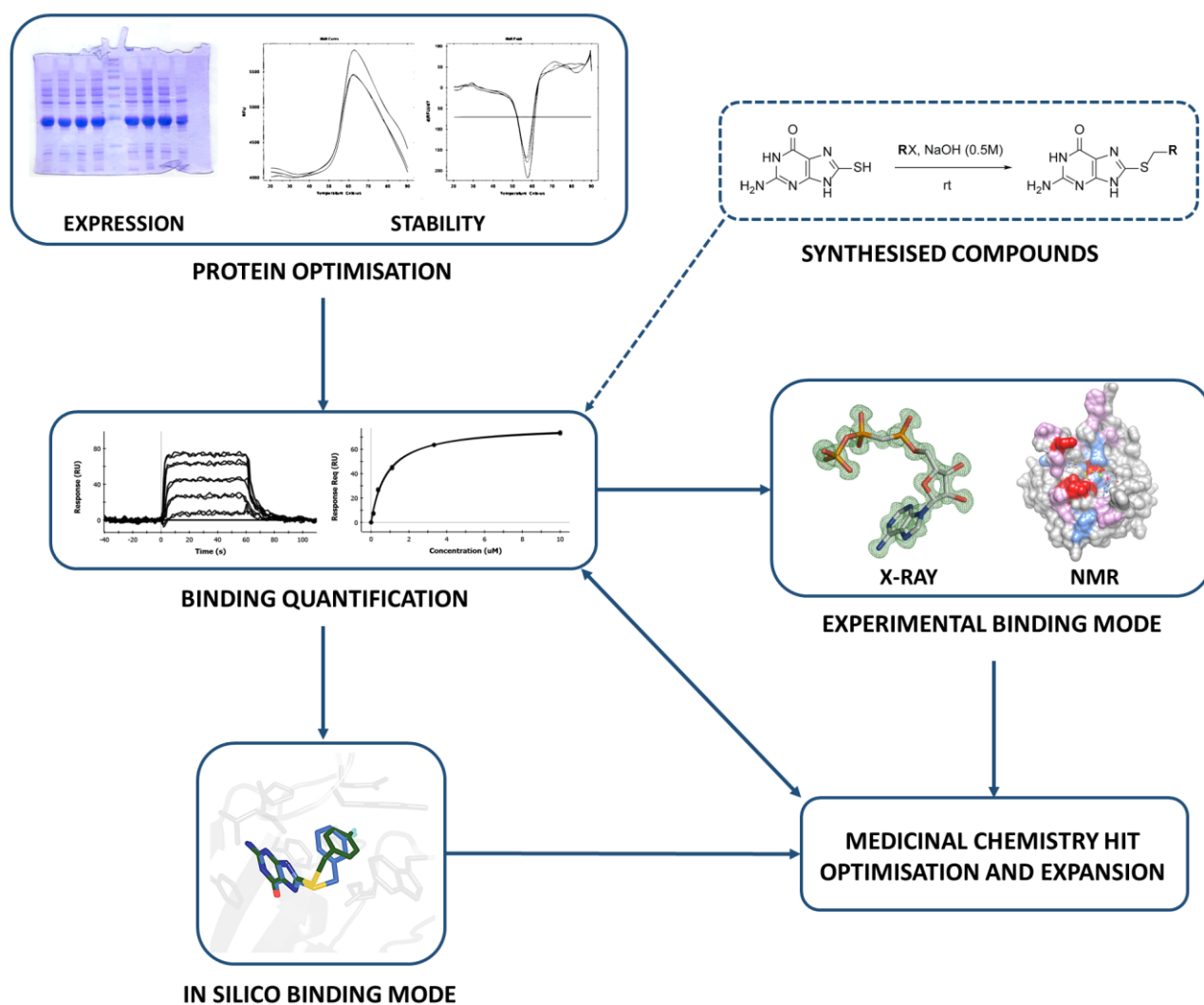


Figure 1.10. General overview of the process employed throughout this project. Synthetic chemistry aspects of this project were performed by collaborators.

1.9 References

1. Bentley, R. Different roads to discovery; Prontosil (hence sulfa drugs) and penicillin (hence β -lactams). *J. Ind. Microbiol. Biotechnol.* **2009**, 36, 775–786.
2. Domagk, G. Twenty-five years of sulfonamide therapy. *Ann. N. Y. Acad. Sci.* **1957**, 69, 380–384.
3. Lesch, J. *The first miracle drugs. How the sulfa drugs transformed medicine*. Oxford University Press, New York, NY: 2007.
4. Armstrong, G. L.; Conn, L. A.; Pinner, R. W. Trends in infectious disease mortality in the united states during the 20th century. *JAMA* **1999**, 281, 61–66.
5. Zaffiri, L.; Gardner, J.; Toledo-Pereyra, L. H. History of antibiotics. From salvarsan to cephalosporins. *J. Invest. Surg.* **2012**, 25, 67–77.
6. Davies, J.; Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **2010**, 74, 417–433.
7. Boucher, H.; Talbot, G.; Bradley, J.; Edwards, J.; Gilbert, D.; Rice, L. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2009**, 48, 1–12.
8. Coates, A.; Hu, Y.; Bax, R.; Page, C. The future challenges facing the development of new antimicrobial drugs. *Nat. Rev. Drug Discov.* **2002**, 1, 895–910.
9. Payne, D.; Gwynn, M.; Holmes, D.; Pompliano, D. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **2007**, 6, 29–40.
10. Sefton, A. M. Mechanisms of antimicrobial resistance. *Drugs* **2002**, 62, 557–566.
11. D'Costa, V. M.; King, C. E.; Kalan, L.; Morar, M.; Sung, W. W. L.; Schwarz, C.; Froese, D.; Zazula, G.; Calmels, F.; Debruyne, R.; Golding, G. B.; Poinar, H. N.; Wright, G. D. Antibiotic resistance is ancient. *Nature* **2011**, 477, 457–461.
12. Alanis, A. J. Resistance to antibiotics: are we in the post-antibiotic era? *Arch. Med. Res.* **2005**, 36, 697–705.
13. Sköld, O. Sulfonamide resistance: mechanisms and trends. *Drug Resist. Updat.* **2000**, 3, 155–160.

-
14. World Health Organization. Antimicrobial resistance: global report on surveillance 2014. **2014**, available from: <http://www.who.int/drugresistance/documents/surveillancereport/en> (accessed September 16, 2015).
 15. Piddock, L. J. The crisis of no new antibiotics--what is the way forward? *Lancet Infect. Dis.* **2012**, 12, 249–253.
 16. Shlaes, D. M. Research and development of antibiotics: the next battleground. *ACS Infect. Dis.* **2015**, 1, 232–233.
 17. Silver, L. L. Challenges of antibacterial discovery. *Clin. Microbiol. Rev.* **2011**, 24, 71–109.
 18. Projan, S. J. Why is big Pharma getting out of antibacterial drug discovery? *Curr. Opin. Microbiol.* **2003**, 6, 427–430.
 19. White, A. R. Effective antibacterials: at what cost? The economics of antibacterial resistance and its control. *J. Antimicrob. Chemother.* **2011**, 66, 1948–1953.
 20. Infectious Diseases Society of America. Combating antimicrobial resistance: policy recommendations to save lives. *Clin. Infect. Dis.* **2011**, 52, S397–S428.
 21. Rex, J. H.; Eisenstein, B. I.; Alder, J.; Goldberger, M.; Meyer, R.; Dane, A.; Friedland, I.; Knirsch, C.; Sanhai, W. R.; Tomayko, J.; Lancaster, C.; Jackson, J. A comprehensive regulatory framework to address the unmet need for new antibacterial treatments. *Lancet Infect. Dis.* **2013**, 13, 269–275.
 22. Laxminarayan, R.; Powers, J. H. Antibacterial R&D incentives. *Nat. Rev. Drug Discov.* **2011**, 10, 727–728.
 23. Fauci, A. S.; Morens, D. M. The perpetual challenge of infectious diseases. *N. Engl. J. Med.* **2012**, 366, 454–461.
 24. Fischbach, M. A.; Walsh, C. T. Antibiotics for emerging pathogens. *Science* **2009**, 325, 1089–1093.
 25. Fisher, J. F.; Mobashery, S. Endless resistance. Endless antibiotics? *MedChemComm* **2016**, 7, 37–49.
 26. Diekema, D. J.; Pfaller, M. A.; Schmitz, F. J.; Smayevsky, J.; Bell, J.; Jones, R. N.; Beach, M. Survey of infections due to *Staphylococcus* species: frequency of occurrence

and antimicrobial susceptibility of isolates collected in the United States, Canada, Latin America, Europe, and the Western Pacific region for the SENTRY Antimicrobial Surveillance Program, 1997-1999. *Clin. Infect. Dis.* **2001**, 32, S114–S132.

27. DeLeo, F. R.; Chambers, H. F. Reemergence of antibiotic-resistant *Staphylococcus aureus* in the genomics era. *J. Clin. Invest.* **2009**, 119, 2464–2474.

28. Kuehnert, M. J.; Kruszon-Moran, D.; Hill, H. A.; McQuillan, G.; McAllister, S. K.; Fosheim, G.; McDougal, L. K.; Chaitram, J.; Jensen, B.; Fridkin, S. K.; Killgore, G.; Tenover, F. C. Prevalence of *Staphylococcus aureus* nasal colonization in the United States, 2001-2002. *J. Infect. Dis.* **2006**, 193, 172–179.

29. Wertheim, H. F. L.; Melles, D. C.; Vos, M. C.; van Leeuwen, W.; van Belkum, A.; Verbrugh, H. A.; Nouwen, J. L. The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect. Dis.* **2005**, 5, 751–762.

30. van Hal, S. J.; Jensen, S. O.; Vaska, V. L.; Espedido, B. A.; Paterson, D. L.; Gosbell, I. B. Predictors of mortality in *Staphylococcus aureus* bacteremia. *Clin. Microbiol. Rev.* **2012**, 25, 362–386.

31. Rountree, P. M.; Freeman, B. M. Infections caused by a particular phage type of *Staphylococcus aureus*. *Med. J. Aust.* **1955**, 42, 156–161.

32. Kirby, W. M. Extraction of a highly potent penicillin inactivator from penicillin resistant *Staphylococci*. *Science* **1944**, 99, 452–453.

33. Jevons, M. P. “Celbenin” - resistant *Staphylococci*. *Br. Med. J.* **1961**, 1, 124–125.

34. Dantes, R.; Mu, Y.; Belflower, R.; Aragon, D.; Dumyati, G.; Harrison, L. H.; Lessa, F. C.; Lynfield, R.; Nadle, J.; Petit, S.; Ray, S. M.; Schaffner, W.; Townes, J.; Fridkin, S. National burden of invasive methicillin-resistant *staphylococcus aureus* infections, united states, 2011. *JAMA Intern. Med.* **2013**, 173, 1970–1978.

35. Klevens, R. M.; Edwards, J. R.; Tenover, F. C.; McDonald, L. C.; Horan, T.; Gaynes, R. Changes in the epidemiology of methicillin-resistant *Staphylococcus aureus* in intensive care units in US hospitals, 1992-2003. *Clin. Infect. Dis.* **2006**, 42, 389–391.

36. Klein, E.; Smith, D. L.; Laxminarayan, R. Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999–2005. *Emerg. Infect. Disease* **2007**, 13, 1840–1846.

37. Jarvis, W. R.; Schlosser, J.; Chinn, R. Y.; Tweeten, S.; Jackson, M. National prevalence of methicillin-resistant *Staphylococcus aureus* in inpatients at US health care facilities, 2006. *Am. J. Infect. Control* **2007**, 35, 631–637.
38. DeLeo, F. R.; Otto, M.; Kreiswirth, B. N.; Chambers, H. F. Community-associated methicillin-resistant *Staphylococcus aureus*. *Lancet* **2010**, 375, 1557–1568.
39. Herold, B. C.; Immergluck, L. C.; Maranan, M. C.; Lauderdale, D. S.; Gaskin, R. E.; Boyle-Vavra, S.; Leitch, C. D.; Daum, R. S. Community-acquired methicillin-resistant *Staphylococcus aureus* in children with no identified predisposing risk. *JAMA* **1998**, 279, 593–598.
40. Tristan, A.; Bes, M.; Meugnier, H.; Lina, G.; Bozdogan, B.; Courvalin, P.; Reverdy, M.-E.; Enright, M. C.; Vandenesch, F.; Etienne, J. Global distribution of Panton-Valentine leukocidin–positive methicillin-resistant *Staphylococcus aureus*, 2006. *Emerg. Infect. Disease* **2007**, 13, 594.
41. Bubeck-Wardenburg, J.; Bae, T.; Otto, M.; DeLeo, F. R.; Schneewind, O. Poring over pores: [alpha]-hemolysin and Panton-Valentine leukocidin in *Staphylococcus aureus* pneumonia. *Nat. Med.* **2007**, 13, 1405–1406.
42. Wang, R.; Braughton, K. R.; Kretschmer, D.; Bach, T.-H. L.; Queck, S. Y.; Li, M.; Kennedy, A. D.; Dorward, D. W.; Klebanoff, S. J.; Peschel, A.; DeLeo, F. R.; Otto, M. Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. *Nat. Med.* **2007**, 13, 1510–1514.
43. Diep, B. A.; Stone, G. G.; Basuino, L.; Graber, C. J.; Miller, A.; des Etages, S.-A.; Jones, A.; Palazzolo-Ballance, A. M.; Perdreau-Remington, F.; Sensabaugh, G. F.; DeLeo, F. R.; Chambers, H. F. The arginine catabolic mobile element and staphylococcal chromosomal cassette mec linkage: convergence of virulence and resistance in the USA300 clone of methicillin-resistant *Staphylococcus aureus*. *J. Infect. Dis.* **2008**, 197, 1523–1530.
44. French, G. L. Bactericidal agents in the treatment of MRSA infections--the potential role of daptomycin. *J. Antimicrob. Chemother.* **2006**, 58, 1107–1117.
45. Chambers, H. F.; Deleo, F. R. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* **2009**, 7, 629–641.

46. Jang, S. Multidrug efflux pumps in *Staphylococcus aureus* and their clinical implications. *J. Microbiol.* **2016**, *54*, 1–8.
47. Rybak, M. J.; Lomaestro, B. M.; Rotschafer, J. C.; Moellering, R. C.; Craig, W. A.; Billeter, M.; Dalovisio, J. R.; Levine, D. P.; Levine, D. P. Vancomycin therapeutic guidelines: a summary of consensus recommendations from the infectious diseases Society of America, the American Society of Health-System Pharmacists, and the Society of Infectious Diseases Pharmacists. *Clin. Infect. Dis.* **2009**, *49*, 325–327.
48. Elyasi, S.; Khalili, H.; Dashti-Khavidaki, S.; Mohammadpour, A. Vancomycin-induced nephrotoxicity: mechanism, incidence, risk factors and special populations. A literature review. *Eur. J. Clin. Pharmacol.* **2012**, *68*, 1243–1255.
49. Hiramatsu, K. Vancomycin-resistant *Staphylococcus aureus*: a new model of antibiotic resistance. *Lancet Infect. Dis.* **2001**, *1*, 147–155.
50. Hiramatsu, K.; Hanaki, H.; Ino, T.; Yabuta, K.; Oguri, T.; Tenover, F. C. Methicillin-resistant *Staphylococcus aureus* clinical strain with reduced vancomycin susceptibility. *J. Antimicrob. Chemother.* **1997**, *40*, 135–136.
51. Melo-Cristino, J.; Resina, C.; Manuel, V.; Lito, L.; Ramirez, M. First case of infection with vancomycin-resistant *Staphylococcus aureus* in Europe. *Lancet* **2013**, *382*, 205.
52. Centers for Disease Control and Prevention (CDC). *Staphylococcus aureus* resistant to vancomycin--United States, 2002. *MMWR Morb. Mortal. Wkly. Rep.* **2002**, *51*, 565–567.
53. Wunderink, R. G.; Niederman, M. S.; Kollef, M. H.; Shorr, A. F.; Kunkel, M. J.; Baruch, A.; McGee, W. T.; Reisman, A.; Chastre, J. Linezolid in methicillin-resistant *Staphylococcus aureus* nosocomial pneumonia: a randomized, controlled study. *Clin. Infect. Dis.* **2012**, *54*, 621–629.
54. Grim, S. A.; Rapp, R. P.; Martin, C. A.; Evans, M. E. Trimethoprim-sulfamethoxazole as a viable treatment option for infections caused by methicillin-resistant *staphylococcus aureus*. *Pharmacotherapy* **2005**, *25*, 253–264.
55. Campbell, M. L.; Marchaim, D.; Pogue, J. M.; Sunkara, B.; Bheemreddy, S.; Bathina, P.; Pulluru, H.; Chugh, N.; Wilson, M. N.; Moshos, J.; Ku, K.; Hayakawa, K.; Martin, E. T.; Lephart, P. R.; Rybak, M. J.; Kaye, K. S. Treatment of methicillin-resistant *Staphylococcus aureus* infections with a minimal inhibitory concentration of 2 µg/mL to vancomycin: old

(trimethoprim/sulfamethoxazole) versus new (daptomycin or linezolid) agents. *Ann. Pharmacother.* **2012**, *46*, 1587–1597.

56. Miller, L. G.; Perdreau-Remington, F.; Rieg, G.; Mehdi, S.; Perlroth, J.; Bayer, A. S.; Tang, A. W.; Phung, T. O.; Spellberg, B. Necrotizing fasciitis caused by community-associated methicillin-resistant *Staphylococcus aureus* in Los Angeles. *N. Engl. J. Med.* **2005**, *352*, 1445–53.

57. Tenaillon, O.; Skurnik, D.; Picard, B.; Denamur, E. The population genetics of commensal *Escherichia coli*. *Nat. Rev. Micro.* **2010**, *8*, 207–217.

58. Hobman, J. L.; Penn, C. W.; Pallen, M. J. Laboratory strains of *Escherichia coli*: model citizens or deceitful delinquents growing old disgracefully? *Mol. Microbiol.* **2007**, *64*, 881–885.

59. Huang, C. J.; Lin, H.; Yang, X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J. Ind. Microbiol. Biotechnol.* **2012**, *39*, 383–399.

60. Gräslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; dhe-Paganon, S.; Park, H.; Savchenko, A.; Yee, A.; Edwards, A.; Vincentelli, R.; Cambillau, C.; Kim, R.; Kim, S.; Rao, Z.; Shi, Y.; Terwilliger, T.; Kim, C.; Hung, L.; Waldo, G.; Peleg, Y.; Albeck, S.; Unger, T.; Dym, O.; Prilusky, J.; Sussman, J.; Stevens, R.; Lesley, S.; Wilson, I.; Joachimiak, A.; Collart, F.; Dementieva, I.; Donnelly, M.; Eschenfeldt, W.; Kim, Y.; Stols, L.; Wu, R.; Zhou, M.; Burley, S.; Emtage, J.; Sauder, J.; Thompson, D.; Bain, K.; Luz, J.; Gheyi, T.; Zhang, F.; Atwell, S.; Almo, S.; Bonanno, J.; Fiser, A.; Swaminathan, S.; Studier, F.; Chance, M.; Sali, A.; Acton, T.; Xiao, R.; Zhao, L.; Ma, L.; Hunt, J.; Tong, L.; Cunningham, K.; Inouye, M.; Anderson, S.; Janjua, H.; Shastry, R.; Ho, C.; Wang, D.; Wang, H.; Jiang, M.; Montelione, G.; Stuart, D.; Owens, R.; Daenke, S.; Schütz, A.; Heinemann, U.; Yokoyama, S.; Büsow, K.; Gunsalus, K. Protein production and purification. *Nat. Meth.* **2008**, *5*, 135–146.

61. Jana, S.; Deb, J. K. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2005**, *67*, 289–298.

62. Peti, W.; Page, R. Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expr. Purif.* **2007**, *51*, 1–10.

63. Villaverde, A.; Carrio, M. M. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol. Lett.* **2003**, *25*, 1385–1395.
64. Lindhout, D. A.; Thiessen, A.; Schieve, D.; Sykes, B. D. High-yield expression of isotopically labeled peptides for use in NMR studies. *Protein Sci.* **2003**, *12*, 1786–1791.
65. Berg, R. D. The indigenous gastrointestinal microflora. *Trends Microbiol.* **1996**, *4*, 430–435.
66. Kaper, J. B.; Nataro, J. P.; Mobley, H. L. T. Pathogenic *Escherichia coli*. *Nat. Rev. Micro.* **2004**, *2*, 123–140.
67. Croxen, M. A.; Law, R. J.; Scholz, R.; Keeney, K. M.; Wlodarska, M.; Finlay, B. B. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin. Microbiol. Rev.* **2013**, *26*, 822–880.
68. Kotloff, K. L.; Nataro, J. P.; Blackwelder, W. C.; Nasrin, D.; Farag, T. H.; Panchalingam, S.; Wu, Y.; Sow, S. O.; Sur, D.; Breiman, R. F.; Faruque, A. S. G.; Zaidi, A. K. M.; Saha, D.; Alonso, P. L.; Tamboura, B.; Sanogo, D.; Onwuchekwa, U.; Manna, B.; Ramamurthy, T.; Kanungo, S.; Ochieng, J. B.; Omore, R.; Oundo, J. O.; Hossain, A.; Das, S. K.; Ahmed, S.; Qureshi, S.; Quadri, F.; Adegbola, R. A.; Antonio, M.; Hossain, M. J.; Akinsola, A.; Mandomando, I.; Nhampossa, T.; Acácio, S.; Biswas, K.; O'Reilly, C. E.; Mintz, E. D.; Berkeley, L. Y.; Muhsen, K.; Sommerfelt, H.; Robins-Browne, R. M.; Levine, M. M. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* **2013**, *382*, 209–222.
69. Nataro, J. P.; Kaper, J. B. Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **1998**, *11*, 142–201.
70. World Health Organization 2012. World health statistics 2012 WHO Press, Geneva, Switzerland.
71. Taylor, D. N.; Bourgeois, A. L.; Ericsson, C. D.; Steffen, R.; Jiang, Z.-D.; Halpern, J.; Haake, R.; Dupont, H. L. A randomized, double-blind, multicenter study of rifaximin compared with placebo and with ciprofloxacin in the treatment of travelers' diarrhea. *Am. J. Trop. Med. Hyg.* **2006**, *74*, 1060–1066.
72. Dupont, H. L.; Jiang, Z. D.; Belkind-Gerson, J.; Okhuysen, P. C.; Ericsson, C. D.; Ke, S.; Huang, D. B.; Dupont, M. W.; Adachi, J. A.; De La Cabada, F. J.; Taylor, D. N.; Jaini, S.;

Martinez Sandoval, F. Treatment of travelers' diarrhea: randomized trial comparing rifaximin, rifaximin plus loperamide, and loperamide alone. *Clin. Gastroenterol. Hepatol.* **2007**, *5*, 451–456.

73. Centers for Disease Control and Prevention (CDC). Discovery of first mcr-1 gene in *E. coli* bacteria found in a human in United States. In 2016.

74. Vollaard, E. J.; Clasener, H. A. Colonization resistance. *Antimicrob. Agents Chemother.* **1994**, *38*, 409–414.

75. Hudault, S.; Guignot, J.; Servin, A. *Escherichia coli* strains colonising the gastrointestinal tract protect germfree mice against *Salmonella typhimurium* infection. *Gut* **2001**, *49*, 47–55.

76. Rastegar Lari, A.; Gold, F.; Borderon, J. C.; Laugier, J.; Lafont, J. P. Implantation and in vivo antagonistic effects of antibiotic-susceptible *Escherichia coli* strains administered to premature newborns. *Biol. Neonate* **1990**, *58*, 73–78.

77. Fowler, S.; Webber, A.; Cooper, B. S.; Phimister, A.; Price, K.; Carter, Y.; Kibbler, C. C.; Simpson, A. J.; Stone, S. P. Successful use of feedback to improve antibiotic prescribing and reduce *Clostridium difficile* infection: a controlled interrupted time series. *J. Antimicrob. Chemother.* **2007**, *59*, 990–995.

78. Lemon, K. P.; Armitage, G. C.; Relman, D. A.; Fischbach, M. A. Microbiota-targeted therapies: an ecological perspective. *Sci. Transl. Med.* **2012**, *4*, 137rv5.

79. Buffie, C. G.; Jarchum, I.; Equinda, M.; Lipuma, L.; Gobourne, A.; Viale, A.; Ubeda, C.; Xavier, J.; Pamer, E. G. Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to *Clostridium difficile*-induced colitis. *Infect. Immun.* **2012**, *80*, 62–73.

80. Yates, C. M.; Shaw, D. J.; Roe, A. J.; Woolhouse, M. E.; Amyes, S. G. Enhancement of bacterial competitive fitness by apramycin resistance plasmids from non-pathogenic *Escherichia coli*. *Biol. Lett.* **2006**, *2*, 463–465.

81. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **2013**, *12*, 371–387.

82. Swarbrick, J.; Iliades, P.; Simpson, J. S.; Macreadie, I. Folate biosynthesis - reappraisal of old and novel targets in the search for new antimicrobials. *Open Enzym. Inhib. J.* **2008**, *1*, 12–33.
83. Bermingham, A.; Derrick, J. P. The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. *Bioessays* **2002**, *24*, 637–648.
84. Jabrin, S.; Ravanel, S.; Gambonnet, B.; Douce, R.; Rébeillé, F. One-carbon metabolism in plants. Regulation of tetrahydrofolate synthesis during germination and seedling development. *Plant Physiol.* **2003**, *131*, 1431–1439.
85. Nzila, A. Inhibitors of de novo folate enzymes in *Plasmodium falciparum*. *Drug Discov. Today* **2006**, *11*, 939–944.
86. Bourne, C. R. Utility of the biosynthetic folate pathway for targets in antimicrobial discovery. *Antibiotics* **2014**, *3*, 1–28.
87. Sincak, C. A.; Schmidt, J. M. Iclaprim, a novel diaminopyrimidine for the treatment of resistant gram-positive infections. *Ann. Pharmacother.* **2009**, *43*, 1107–1114.
88. Keiser, M. J.; Roth, B. L.; Armbruster, B. N.; Ernsberger, P.; Irwin, J. J.; Shoichet, B. K. Relating protein pharmacology by ligand chemistry. *Nat. Biotech.* **2007**, *25*, 197–206.
89. Pemble, C. W. t.; Mehta, P. K.; Mehra, S.; Li, Z.; Nourse, A.; Lee, R. E.; White, S. W. Crystal structure of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase* dihydropteroate synthase bifunctional enzyme from *Francisella tularensis*. *PLoS One* **2010**, *5*, e14165.
90. Chaudhuri, R. R.; Allen, A. G.; Owen, P. J.; Shalom, G.; Stone, K.; Harrison, M.; Burgis, T. A.; Lockyer, M.; Garcia-Lara, J.; Foster, S. J.; Pleasance, S. J.; Peters, S. E.; Maskell, D. J.; Charles, I. G. Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **2009**, *10*, 291.
91. Blaszczyk, J.; Shi, G.; Li, Y.; Yan, H.; Ji, X. Reaction trajectory of pyrophosphoryl transfer catalyzed by 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Structure* **2004**, *12*, 467–475.
92. Derrick, J. P. The structure and mechanism of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Vitam. Horm.* **2008**, *79*, 411–433.

93. Xiao, B.; Shi, G.; Chen, X.; Yan, H.; Ji, X. Crystal structure of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, a potential target for the development of novel antimicrobial agents. *Structure* **1999**, *7*, 489–496.
94. Hennig, M.; Dale, G. E.; D'arcy, A.; Danel, F.; Fischer, S.; Gray, C. P.; Jolidon, S.; Müller, F.; Page, M. G.; Pattison, P.; Oefner, C. The structure and function of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Haemophilus influenzae*. *J. Mol. Biol.* **1999**, *287*, 211–219.
95. Garçon, A.; Levy, C.; Derrick, J. P. Crystal structure of the bifunctional dihydroneopterin aldolase/6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Streptococcus pneumoniae*. *J. Mol. Biol.* **2006**, *360*, 644–653.
96. Lawrence, M. C.; Iliades, P.; Fernley, R. T.; Berglez, J.; Pilling, P. A.; Macreadie, I. G. The three-dimensional structure of the bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/dihydropteroate synthase of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **2005**, *348*, 655–670.
97. Shaw, G. X.; Li, Y.; Shi, G.; Wu, Y.; Cherry, S.; Needle, D.; Zhang, D.; Tropea, J. E.; Waugh, D. S.; Yan, H.; Ji, X. Structural enzymology and inhibition of the bi-functional folate pathway enzyme HPPK-DHPS from the biowarfare agent *Francisella tularensis*. *FEBS J.* **2014**, *281*, 4123–4137.
98. Chhabra, S.; Dolezal, O.; Collins, B. M.; Newman, J.; Simpson, J. S.; Macreadie, I. G.; Fernley, R.; Peat, T. S.; Swarbrick, J. D. Structure of *S. aureus* HPPK and the discovery of a new substrate site inhibitor. *PLoS One* **2012**, *7*, e29444.
99. Xiao, B.; Shi, G.; Gao, J.; Blaszczyk, J.; Liu, Q.; Ji, X.; Yan, H. Unusual conformational changes in 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase as revealed by X-ray crystallography and NMR. *J. Biol. Chem.* **2001**, *276*, 40274–40281.
100. Lescop, E.; Lu, Z.; Liu, Q.; Xu, H.; Li, G.; Xia, B.; Yan, H.; Jin, C. Dynamics of the conformational transitions in the assembling of the Michaelis complex of a bisubstrate enzyme: a ¹⁵N relaxation study of *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Biochemistry* **2009**, *48*, 302–312.
101. Garçon, A.; Bermingham, A.; Lian, L.-Y.; Derrick, J. P. Kinetic and structural characterization of a product complex of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Escherichia coli*. *Biochem. J.* **2004**, *380*, 867–873.

102. Li, Y.; Gong, Y.; Shi, G.; Blaszczyk, J.; Ji, X.; Yan, H. Chemical transformation is not rate-limiting in the reaction catalyzed by *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Biochemistry* **2002**, *41*, 8777–8783.
103. Li, G.; Felczak, K.; Shi, G.; Yan, H. Mechanism of the conformational transitions in 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase as revealed by NMR spectroscopy. *Biochemistry* **2006**, *45*, 12573–12581.
104. Blaszczyk, J.; Shi, G.; Yan, H.; Ji, X. Catalytic center assembly of HPPK as revealed by the crystal structure of a ternary complex at 1.25 Å resolution. *Structure* **2000**, *8*, 1049–1058.
105. Blaszczyk, J.; Li, Y.; Cherry, S.; Alexandratos, J.; Wu, Y.; Shaw, G.; Tropea, J. E.; Waugh, D. S.; Yan, H.; Ji, X. Structure and activity of *Yersinia pestis* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase as a novel target for the development of antiplague therapeutics. *Acta Crystallogr. D Biol. Crystallogr.* **2007**, *63*, 1169–1177.
106. Bermingham, A.; Bottomley, J. R.; Primrose, W. U.; Derrick, J. P. Equilibrium and kinetic studies of substrate binding to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Escherichia coli*. *J. Biol. Chem.* **2000**, *275*, 17962–17967.
107. Shi, G.; Gong, Y.; Savchenko, A.; Zeikus, J. G.; Xiao, B.; Ji, X.; Yan, H. Dissecting the nucleotide binding properties of *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase with fluorescent 3'(2)'-o-anthraniloyladenine 5'-triphosphate. *Biochim. Biophys. Acta.* **2000**, *1478*, 289–299.
108. Talarico, T. L.; Dev, I. K.; Dallas, W. S.; Ferone, R.; Ray, P. H. Purification and partial characterization of 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase and 7,8-dihydropteroate synthase from *Escherichia coli* MC4100. *J. Bacteriol.* **1991**, *173*, 7029–7032.
109. Stammers, D. K.; Achari, A.; Somers, D. O.; Bryant, P. K.; Rosemond, J.; Scott, D. L.; Champness, J. N. 2.0 Å X-ray structure of the ternary complex of 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase from *Escherichia coli* with ATP and a substrate analogue. *FEBS Lett.* **1999**, *456*, 49–53.
110. Wood, H. Specific inhibition of dihydrofolate biosynthesis - a new approach to chemotherapy. In *Chemistry and Biology of Pteridines*, Walter de Gruyter: Berlin-New York, 1975; pp 27–49.

111. Shi, G.; Blaszczyk, J.; Ji, X.; Yan, H. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: synthesis and biochemical and crystallographic studies. *J. Med. Chem.* **2001**, *44*, 1364–1371.
112. Shi, G.; Shaw, G.; Liang, Y. H.; Subburaman, P.; Li, Y.; Wu, Y.; Yan, H.; Ji, X. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New design with improved properties. *Bioorg. Med. Chem.* **2012**, *20*, 47–57.
113. Shi, G.; Shaw, G.; Li, Y.; Wu, Y.; Yan, H.; Ji, X. Bisubstrate analog inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: new lead exhibits a distinct binding mode. *Bioorg. Med. Chem.* **2012**, *20*, 4303–4309.
114. Shi, G.; Shaw, G.; Ji, X. Linked purine pterin HPPK inhibitors useful as antibacterial agents. US Patent 9,029,344, 2015.
115. Chhabra, S.; Barlow, N.; Dolezal, O.; Hattarki, M. K.; Newman, J.; Peat, T. S.; Graham, B.; Swarbrick, J. D. Exploring the chemical space around 8-mercaptoguanine as a route to new inhibitors of the folate biosynthesis enzyme HPPK. *PLoS One* **2013**, *8*, e59535.
116. Hampele, I. C.; D'Arcy, A.; Dale, G. E.; Kostrewa, D.; Nielsen, J.; Oefner, C.; Page, M. G.; Schönfeld, H. J.; Stüber, D.; Then, R. L. Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. *J. Mol. Biol.* **1997**, *268*, 21–30.
117. Babaoglu, K.; Qi, J.; Lee, R. E.; White, S. W. Crystal structure of 7,8-dihydropteroate synthase from *Bacillus anthracis*: mechanism and novel inhibitor design. *Structure* **2004**, *12*, 1705–1717.
118. Achari, A.; Somers, D. O.; Champness, J. N.; Bryant, P. K.; Rosemond, J.; Stammers, D. K. Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. *Nat. Struct. Biol.* **1997**, *4*, 490–497.
119. Yun, M. K.; Wu, Y.; Li, Z.; Zhao, Y.; Waddell, M. B.; Ferreira, A. M.; Lee, R. E.; Bashford, D.; White, S. W. Catalysis and sulfa drug resistance in dihydropteroate synthase. *Science* **2012**, *335*, 1110–1114.
120. Hammoudeh, D. I.; Zhao, Y.; White, S. W.; Lee, R. E. Replacing sulfa drugs with novel DHPS inhibitors. *Future Med. Chem.* **2013**, *5*, 1331–40.

121. Hevener, K. E.; Yun, M. K.; Qi, J.; Kerr, I. D.; Babaoglu, K.; Hurdle, J. G.; Balakrishna, K.; White, S. W.; Lee, R. E. Structural studies of pterin-based inhibitors of dihydropteroate synthase. *J. Med. Chem.* **2010**, *53*, 166–177.
122. Zhao, Y.; Hammoudeh, D.; Yun, M. K.; Qi, J.; White, S. W.; Lee, R. E. Structure-based design of novel pyrimido[4,5-c]pyridazine derivatives as dihydropteroate synthase inhibitors with increased affinity. *ChemMedChem* **2012**, *7*, 861–870.
123. Morrison, R. W.; Mallory, W. R.; Styles, V. L. Pyrimido[4,5-c]pyridazines. 1. Cyclizations with .alpha.-keto esters. *J. Org. Chem.* **1978**, *43*, 4844–4849.
124. Morrison, R., Jr; Mallory, W.; Styles, V. EP 0000383 (A1) Pyrimido[4,5-c]pyridazines, their use in pharmaceutical compositions, process and intermediates for their preparation. 1979.
125. Lever, O. J.; Bell, L.; McGuire, H.; Ferone, R. Monocyclic pteridine analogues. Inhibition of *Escherichia coli* dihydropteroate synthase by 6-amino-5-nitrosoisocytosines. *J. Med. Chem.* **1985**, *28*, 1870–1874.
126. Lever, O. W. J.; Bell, L. N.; Hyman, C.; McGuire, H. M.; Ferone, R. Inhibitors of dihydropteroate synthase: substituent effects in the side-chain aromatic ring of 6-[[3-(aryloxy)propyl]amino]-5-nitrosoisocytosines and synthesis and inhibitory potency of bridged 5-nitrosoisocytosine-p-aminobenzoic acid analogues. *J. Med. Chem.* **1986**, *29*, 665–670.
127. Zhao, Y.; Shadrack, W. R.; Wallace, M. J.; Wu, Y.; Griffith, E. C.; Qi, J.; Yun, M.-K.; White, S. W.; Lee, R. E. Pterin–sulfa conjugates as dihydropteroate synthase inhibitors and antibacterial agents. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 3950–4.
128. Qi, J.; Virga, K.; Das, S.; Zhao, Y.; Yun, M.; White, S.; Lee, R. L. Synthesis of bi-substrate state mimics of dihydropteroate synthase as potential inhibitors and molecular probes. *Bioorg Med Chem* **2011**, *19*, 1298–305.
129. Hammoudeh, D. I.; Daté, M.; Yun, M. K.; Zhang, W.; Boyd, V. A.; Viacava Follis, A.; Griffith, E.; Lee, R. E.; Bashford, D.; White, S. W. Identification and characterization of an allosteric inhibitory site on dihydropteroate synthase. *ACS Chem. Biol.* **2014**, *9*, 1294–1302.
130. Bleicher, K. H.; Bohm, H.-J.; Muller, K.; Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nat. Rev. Drug Discov.* **2003**, *2*, 369–378.

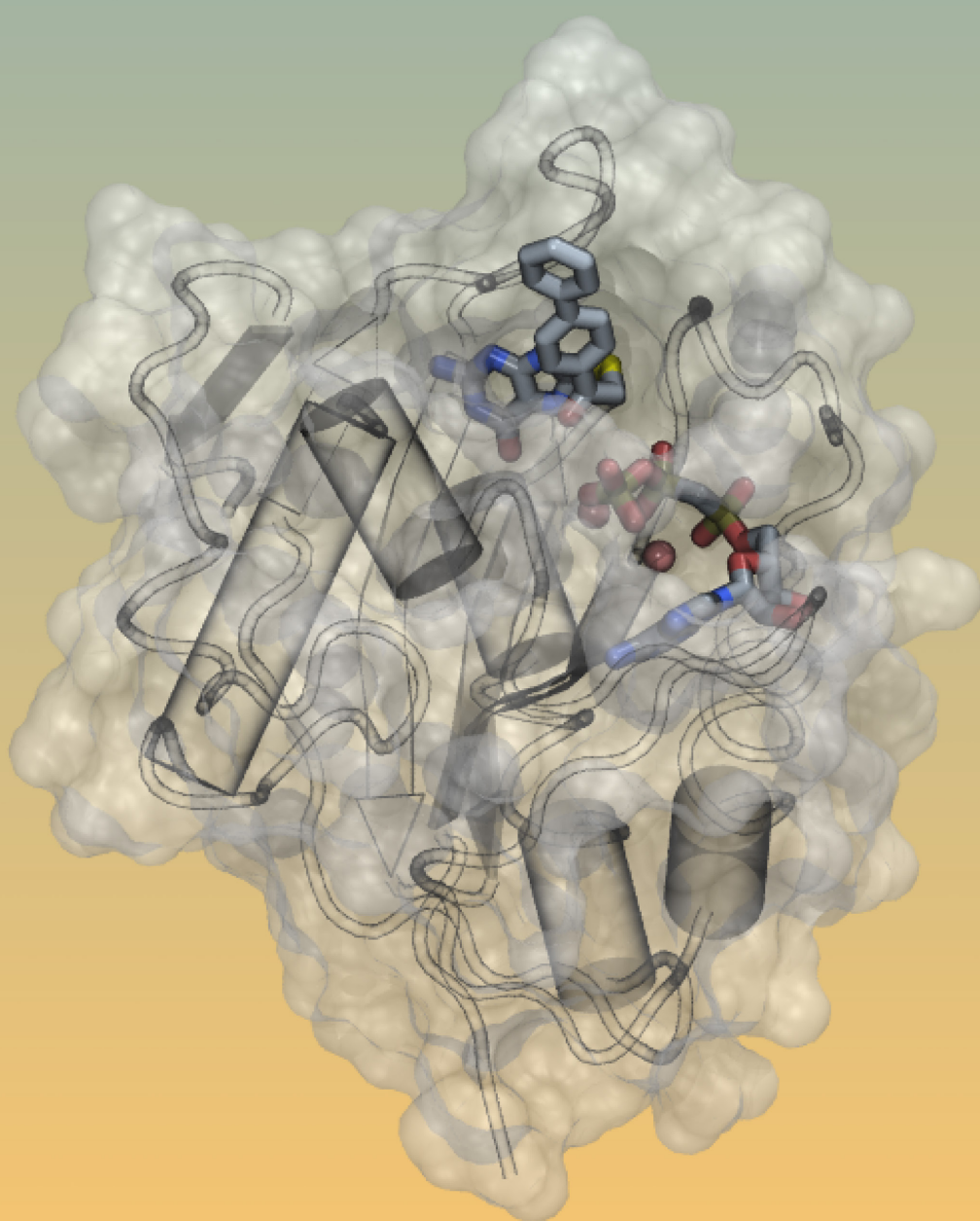
131. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.
132. Veber, D.; Johnson, S.; Cheng, H.; Smith, B.; Ward, K.; Kopple, K. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615–23.
133. Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* **2004**, *3*, 711–716.
134. Hann, M. M.; Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255–263.
135. Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S. Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **2011**, *10*, 188–195.
136. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.
137. Tsai, J.; Lee, J. T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N. K.; Sproesser, K.; Li, L.; Smalley, K. S. M.; Fong, D.; Zhu, Y.-L.; Marimuthu, A.; Nguyen, H.; Lam, B.; Liu, J.; Cheung, I.; Rice, J.; Suzuki, Y.; Luu, C.; Settachatgul, C.; Shellooe, R.; Cantwell, J.; Kim, S.-H.; Schlessinger, J.; Zhang, K. Y. J.; West, B. L.; Powell, B.; Habets, G.; Zhang, C.; Ibrahim, P. N.; Hirth, P.; Artis, D. R.; Herlyn, M.; Bollag, G. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *PNAS* **2008**, *105*, 3041–3046.
138. Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; Burton, E. A.; Wong, B.; Tsang, G.; West, B. L.; Powell, B.; Shellooe, R.; Marimuthu, A.; Nguyen, H.; Zhang, K. Y. J.; Artis, D. R.; Schlessinger, J.; Su, F.; Higgins, B.; Iyer, R.; D'Andrea, K.; Koehler, A.; Stumm, M.; Lin, P. S.; Lee, R. J.; Grippo, J.; Puzanov, I.; Kim, K. B.; Ribas, A.; McArthur, G. A.; Sosman, J. A.; Chapman, P. B.; Flaherty, K. T.; Xu, X.; Nathanson, K. L.; Nolop, K. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **2010**, *467*, 596–599.

139. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discovery Today* **2003**, *8*, 876–877.
140. Jhoti, H.; Williams, G.; Rees, D. C.; Murray, C. W. The 'rule of three' for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discov.* **2013**, *12*, 644–645.
141. Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **2004**, *432*, 855–861.
142. Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **2012**, *51*, 4990–5003.
143. Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430–431.
144. Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences* **2001**, *41*, 856–864.
145. Leach, A. R.; Hann, M. M. Molecular complexity and fragment-based drug discovery: ten years on. *Current Opinion in Chemical Biology* **2011**, *15*, 489–496.
146. Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308–1315.
147. Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007**, *2*, 2212–21.
148. Zartler, E. R.; Shapiro, M. J. Protein NMR-based screening in drug discovery. *Curr. Pharm. Des.* **2006**, *12*, 3963–3972.
149. Caliandro, R.; Belviso, D. B.; Aresta, B. M.; de Candia, M.; Altomare, C. D. Protein crystallography and fragment-based drug design. *Future Med. Chem.* **2013**, *5*, 1121–1140.
150. Chavanieu, A.; Pugnère, M. Developments in SPR fragment screening. *Expert Opin. Drug Discov.* **2016**, *11*, 489–499.
151. Edink, E.; Jansen, C.; Leurs, R.; de Esch, I. J. P. The heat is on: thermodynamic analysis in fragment-based drug discovery. *Drug Discov. Today Technol.* **2010**, *7*, e189–e201.

152. Howard, N.; Abell, C.; Blakemore, W.; Chessari, G.; Congreve, M.; Howard, S.; Jhoti, H.; Murray, C. W.; Seavers, L. C. A.; van Montfort, R. L. M. Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. *J. Med. Chem.* **2006**, *49*, 1346–1355.
153. Murray, C. W.; Verdonk, M. L.; Rees, D. C. Experiences in fragment-based drug discovery. *Trends Pharmacol. Sci.* **2012**, *33*, 224–232.
154. Leeson, P. D.; St-Gallay, S. A. The influence of the 'organizational factor' on compound quality in drug discovery. *Nat. Rev. Drug Discov.* **2011**, *10*, 749–765.
155. Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* **2007**, *6*, 211–219.
156. Bai, X.-C.; McMullan, G.; Scheres, S. H. W. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **2015**, *40*, 49–57.

"Structure-based design and development of functionalized mercaptoguanine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*"

2



The following paper continues the work described by Chhabra et al.¹⁻³ in developing the hit fragment, 8MG, into a lead compound. Initially, a commercially available acetophenone derivative of 8MG (**3**, Figure 2.1) is found to improve binding 10-fold relative to the parent compound. A crystal complex of **3** bound to AMPCPP/SaHPPK is pursued and subsequently determined, representing the first solved ternary crystal structure of SaHPPK. These results spur a round of hit progression, in which a series of analogues of **3** are synthesised and tested. A subset of these are reported in this paper, one of which (**7**) is successfully crystallised in complex with AMPCPP/SaHPPK. The lead compound, **11**, improves upon the binding and physicochemical properties of **3**; it represents the tightest binding HPPK inhibitor reported to date (for any species). The binary complex of AMPCPP/SaHPPK is additionally determined and reported. Though not directly advancing the drug discovery aspect of this project, this structure elucidates the cofactor-protein interactions and conformational changes to SaHPPK that occur upon ligand binding.

In a drug discovery setting, NMR techniques will provide data of varying degrees of quality, dependent on factors such as the size, mobility, and stability of the protein being studied. Complementary experiments can be performed to improve the reliability of the data obtained, and approaches as straightforward as increasing the length of an NMR experiment can be implemented to improve data quality. In contrast, crystallographic studies are essentially an “all-or-nothing” affair, since the formation and quality of crystals is dependent on stochastic factors and unpredictable protein-protein interactions; crystallisation of a particular protein or protein-complex may turn out to be inherently intractable.

X-ray crystallography remains the foremost technique for revealing atomic-level information about protein-ligand interactions and the majority of the efforts devoted to the following paper (and indeed, the overall thesis) related to crystallising SaHPPK. At the commencement of this project, it was unknown whether the desired SaHPPK complexes would be crystallisable, and as such, NMR-based efforts were implemented alongside the crystallographic ones. The expression and purification of SaHPPK with isotopic ¹⁵N- and ¹⁵N¹³C-labelling were thus performed, with subsequent backbone NMR assignments of SaHPPK complexed with AMPCPP and either compound **3** or **11** (Figure 2.2). For crystallisation experiments, a thrombin cleavage of the hexahistidine tag of SaHPPK was performed to reduce protein mobility and promote stable crystal contacts. Crystallisation trials were set up using the conditions that had proved successful for an 8MG/SaHPPK complex,¹ however no crystal growth was observed. A rather broad and comprehensive

crystallisation screen (eight 96-well plates) prepared by the CSIRO Collaborative Crystallisation Centre (C3) was implemented with both AMPCPP and **3**/AMPCPP. Fortunately, crystals were obtained for both complexes (under different conditions); optimisation of both sets of crystallisation conditions improved the diffraction obtained, and the two structures were solved. The structure of **7**/AMPCPP/SaHPPK was later also acquired, and used the same conditions determined from the **3**/AMPCPP/SaHPPK trials. This crystal form is very sensitive to the ligand in the complex; crystallisation trials with the analogous compounds **4** and **9**, employing the same conditions, did not result in crystal growth, nor did the same trials with compounds **10** and **11**. The above broad crystallisation screen was furthermore performed with compound **11** and likewise failed to produce crystals.

SPR techniques were used to quantify ligand binding affinities. While ITC experiments were also performed on several of the reported compounds, these experiments were discontinued — the main factor in this decision being the markedly higher protein consumption required. ITC, while providing useful information regarding the thermodynamics of binding, is better suited to proteins expressing with higher yields. This paper focuses on compound binding to SaHPPK, however testing was also performed using the EcDHPS enzyme, the potential for parallel screening representing another benefit of SPR over ITC. This revealed lead compounds for both SaHPPK (presented in the following paper) and EcDHPS (presented in Chapter 4).

Readers should be aware when considering the following paper: although many compounds (particularly the lead compounds 8MG, **3**, and **4**) were re-tested via SPR throughout this thesis, the uncertainties presented for K_D values indicate the standard deviation of technical replicates (i.e. discrete compound sample preparation, but performed on the same chip and therefore not independent experiments), and should be interpreted appropriately.

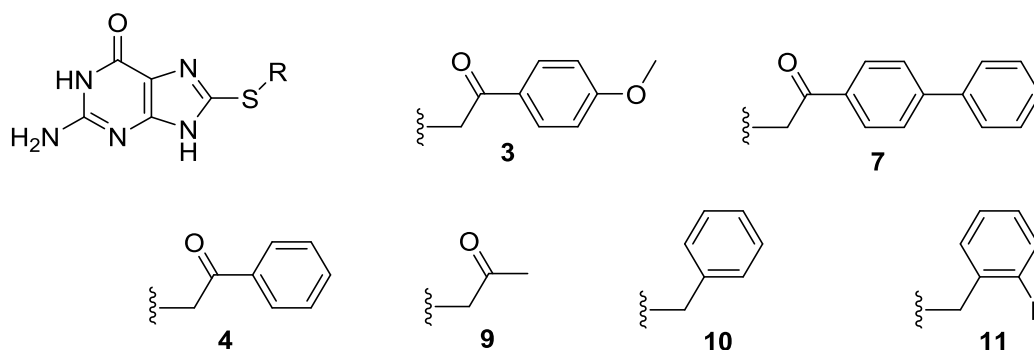


Figure 2.1. Chemical structures of selected compounds reported in the following paper.

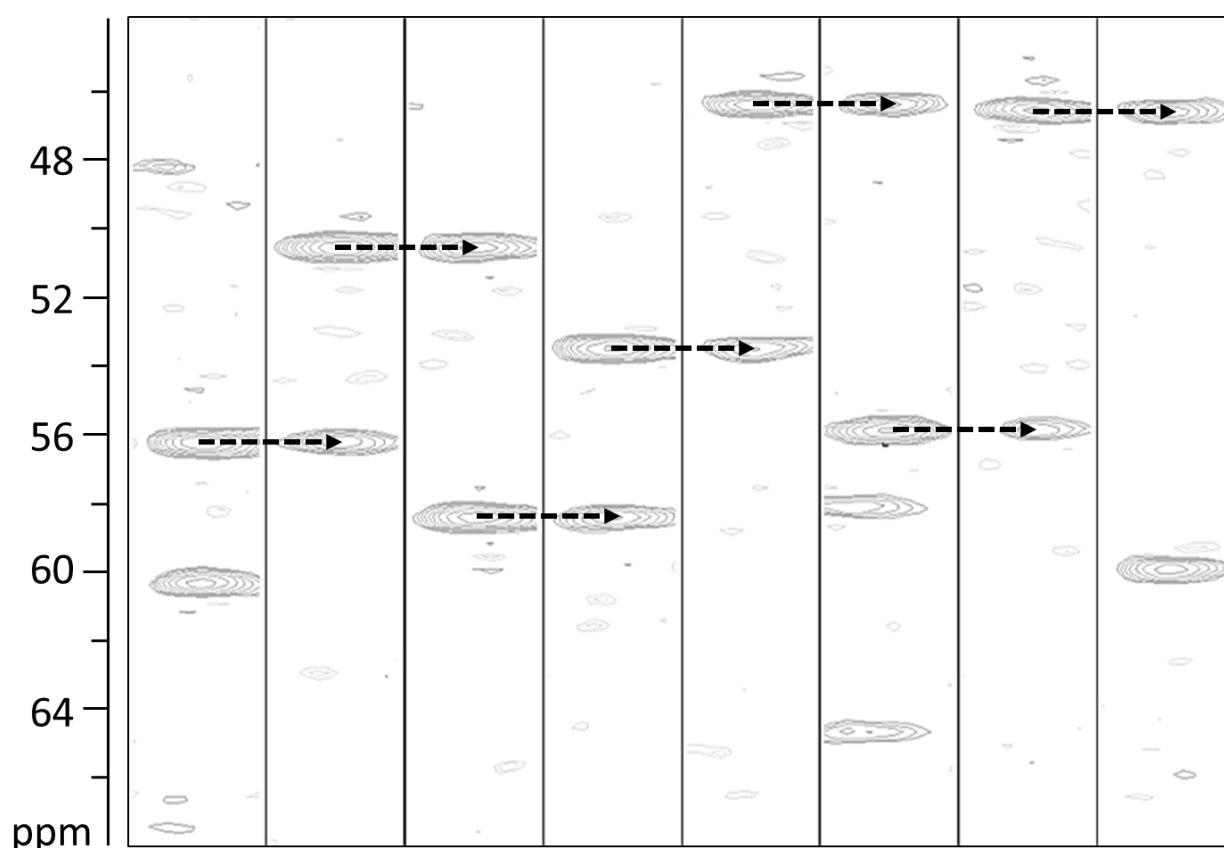


Figure 2.2. Strips of residues 3–10 from the acquired 3D HNCA experiment of the 3/AMPCPP/SaHPPK complex. In each NH strip two $C\alpha$ peaks should be visible; the stronger peak belongs to the same residue, $C\alpha(i)$, and the weaker peak belongs to the preceding residue, $C\alpha(i-1)$. Sequential assignment can be performed by connecting peaks as indicated by the arrows.

1. Chhabra, S.; Dolezal, O.; Collins, B. M.; Newman, J.; Simpson, J. S.; Macreadie, I. G.; Fernley, R.; Peat, T. S.; Swarbrick, J. D. Structure of *S. aureus* HPPK and the discovery of a new substrate site inhibitor. *PLoS One* **2012**, 7, e29444.
2. Chhabra, S.; Barlow, N.; Dolezal, O.; Hattarki, M. K.; Newman, J.; Peat, T. S.; Graham, B.; Swarbrick, J. D. Exploring the chemical space around 8-mercaptopguanine as a route to new inhibitors of the folate biosynthesis enzyme HPPK. *PLoS One* **2013**, 8, e59535.
3. Chhabra, S. Structure and inhibitors of *S. aureus* HPPK - a folate pathway inhibitor. Monash University, Monash Institute of Pharmaceutical Sciences, 2012.

Declaration for Thesis Chapter 2

Chapter 2 presents the published article co-authored with Sandeep Chhabra, Zhong-Chang Wang, Aaron DeBono, Olan Dolezal, Janet Newman, Noel P. Pitcher, Raphaël Rahmani, Ben Cleary, Nicholas Barlow, Meghan Hattarki, Bim Graham, Thomas S. Peat, Jonathan B. Baell, and James D. Swarbrick.

Declaration by candidate

In the case of Chapter 2, the nature and extent of my contribution to the work was as the following:

Nature of Contribution	Extent of Contribution (%)
Protein expression and purification, NMR and X-ray crystallography experiments, SPR, biochemical assays, molecular modelling	70

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Sandeep Chhabra ^{1,2}	Protein expression, intellectual input	
Zhong-Chang Wang ³	Synthetic chemistry	
Aaron DeBono ¹	Synthetic chemistry, intellectual input	
Olan Dolezal ²	SPR experiments, intellectual input	
Janet Newman ²	Crystallisation experiments, intellectual input	
Noel P. Pitcher ¹	Synthetic chemistry	5

Raphaël Rahmani ¹	Supervision of chemists	
Ben Cleary ¹	Supervision of chemists	5
Nicholas Barlow ¹	Supervision of chemists	
Meghan Hattarki ²	SPR experiments	
Bim Graham ¹	Intellectual input	
Thomas S. Peat ²	X-ray crystallography, intellectual input	
Jonathan B. Baell ¹	Synthetic chemistry, intellectual input	
James D. Swarbrick ¹	NMR experiments, intellectual input	

Declaration by co-authors

The undersigned hereby certify that:

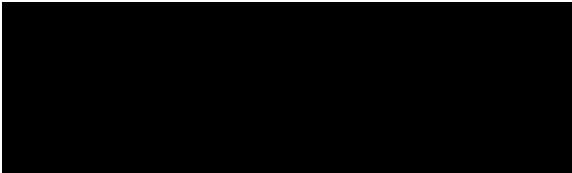
- (1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors;
- (2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- (3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- (4) there are no other authors of the publication according to these criteria;
- (5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- (6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:



Location(s)

¹Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

²CSIRO Biosciences Program, Parkville, Victoria 3052, Australia

³State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, People's Republic of China

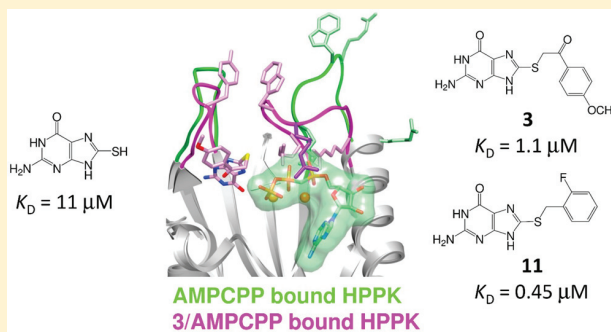
		Date
Signature 1		19/12/2016
Signature 2		
Signature 3		
Signature 4		

		Date
Signature 2		19/12/16
Signature 3		19/12/16
Signature 4		20/12/16

Structure-Based Design and Development of Functionalized Mercaptoguanine Derivatives as Inhibitors of the Folate Biosynthesis Pathway Enzyme 6-Hydroxymethyl-7,8-dihydropterin Pyrophosphokinase from *Staphylococcus aureus*Matthew L. Dennis,^{†,‡} Sandeep Chhabra,^{†,‡} Zhong-Chang Wang,^{†,§} Aaron Debono,[†] Olan Dolezal,[‡] Janet Newman,[‡] Noel P. Pitcher,[†] Raphael Rahmani,[†] Ben Cleary,[†] Nicholas Barlow,[†] Meghan Hattarki,[‡] Bim Graham,[†] Thomas S. Peat,[‡] Jonathan B. Baell,[†] and James D. Swarbrick^{*,†}[†]Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia[‡]CSIRO Biosciences Program, Parkville, Victoria 3052, Australia[§]State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, People's Republic of China

S Supporting Information

ABSTRACT: 6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), an enzyme from the folate biosynthesis pathway, catalyzes the pyrophosphoryl transfer from ATP to 6-hydroxymethyl-7,8-dihydropterin and is a yet-to-be-drugged antimicrobial target. Building on our previous discovery that 8-mercaptopguanine (8MG) is an inhibitor of *Staphylococcus aureus* HPPK (SaHPPK), we have identified and characterized the binding of an S8-functionalized derivative (**3**). X-ray structures of both the SaHPPK/**3**/cofactor analogue ternary and the SaHPPK/cofactor analogue binary complexes have provided insight into cofactor recognition and key residues that move over 30 Å upon binding of **3**, whereas NMR measurements reveal a partially plastic ternary complex active site. Synthesis and binding analysis of a set of analogues of **3** have identified an advanced new lead compound (**11**) displaying >20-fold higher affinity for SaHPPK than 8MG. A number of these exhibited low micromolar affinity for dihydropteroate synthase (DHPS), the adjacent, downstream enzyme to HPPK, and may thus represent promising new leads to bienzyme inhibitors.



■ INTRODUCTION

Folate is an essential vitamin for the growth of all living organisms. The reduced form, tetrahydrofolate, is a critical cofactor for one-carbon transfer reactions required for the synthesis of purines, amino acids, S-adenosylmethionine, thymidine monophosphate, and formyl-methionine.^{1,2} Mammals and higher eukaryotes depend on dietary folate, whereas plants and most microorganisms synthesize it de novo. The sulfa drugs, targeting the folate pathway enzyme dihydropteroate synthase (DHPS), are still used in the clinic after several decades and vindicate the pathway as an attractive source of antimicrobial targets.³ Even today, sulfonamides are coadministered with a bacterial dihydrofolatereductase (DHFR) inhibitor as a synergistic broad-spectrum cocktail to prevent or treat a range of diseases and infections, including malaria, *Toxoplasma gondii* encephalitis, *Pneumocystis carinii* pneumonia, Tuberculosis, and *Staphylococcus aureus* infections.^{4–8}

Point mutations in both DHPS⁹ and DHFR¹⁰ genes have led to resistance to sulfamethoxazole (SMX)- and trimethoprim (TMP)-based therapies, respectively. Structure-based rational

design has therefore sought new generations of lead compounds effective on DHFR mutant strains¹¹ as well as the possibility of more tailored sulfa drugs,¹² novel pterin-site binding motifs,^{9,13} and allosteric inhibitors of DHPS.¹⁴ Recent insights into the structural basis for the off-target side effects of the sulfa drugs¹⁵ are likely to assist in improving the selectivity and efficacy of this class of drug.

Notwithstanding these efforts, the magnitude of the problem of antibiotic resistance is highlighted by the emergence of the methicillin-resistant strains of *S. aureus* (MRSA) (“the super bug”). Originally confined within the hospital setting, more recently it has spread to the community and begun affecting those without any risk factors.^{16,17} The search for effective future treatments has resulted in increasing interest into alternative enzyme targets for new antibacterials¹⁸ as well as novel immunization strategies.¹⁹ In this regard, our own research^{20,21} has focused on the structure-based development of inhibitors of

Received: August 21, 2014

Published: October 30, 2014

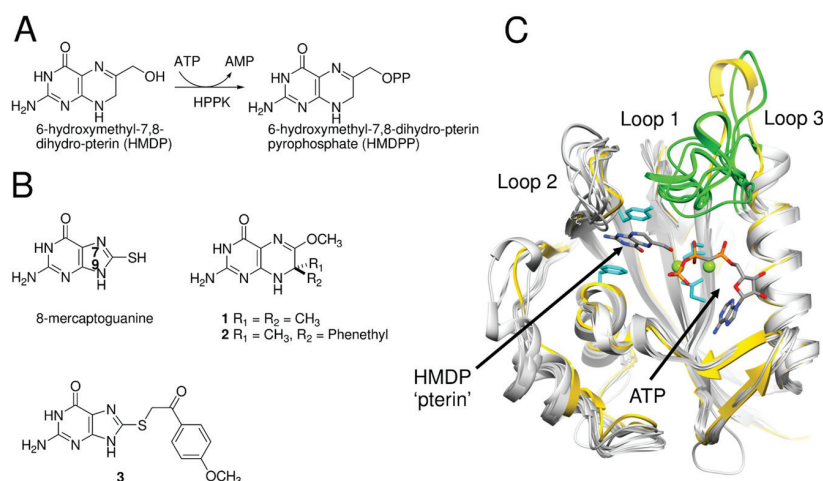


Figure 1. (A) HPPK catalysis. (B) Structures of selected HPPK inhibitors. (C) Superposition of EcHPPK structures (gray) with the SaHPPK/8MG structure (PDB: 3QBC, yellow) showing active site loop conformations in response to a variety of bound ligands along with the bound substrate (HMDP) and cofactor analogue (AMPCPP). Observed loop 3 conformational changes for EcHPPK are highlighted in green.

6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), the enzyme directly preceding DHPS in the folate pathway, responsible for catalyzing pyrophosphoryl transfer from a magnesium-bound ATP cofactor to 6-hydroxymethyl-7,8-dihydro-pterin (HMDP) (Figure 1A). HPPK is not a known target for any existing antibiotic, has no relevant human homologues, and has druggable sites to develop antimicrobials effective on current and future SMX/TMP-resistant pathogens.

HPPK is an 18 kDa enzyme composed of a three-layered α - β - α , thioredoxin-like fold (Figure 1).²² Catalysis follows an ordered mechanism, whereby the ATP cofactor binds initially to a comparatively open pocket in which the triphosphate is complexed by two magnesium ions that are additionally coordinated by two highly conserved aspartic acid residues.²³ Binding of the HMDP substrate then follows, stacking between two conserved aromatic rings. Three loops (loops 1–3, Figure 1) interact and seal the active site to allow pyrophosphate transfer from the cofactor to the substrate. A large body of structural data,²² predominantly for the *Escherichia coli* enzyme (EcHPPK) with the nonhydrolyzable cofactor analogue, AMPCPP, reveals a relatively rigid core structure and flexible loops, particularly the cofactor loop 3, which undergoes a mechanistically important 20 Å positional change during the catalytic cycle.^{23,24} X-ray structures of HPPK from several species^{21,25–28} reveal highly conserved residues in the active site, indicating that suitably designed inhibitors could display broad-spectrum activity.

Only a few inhibitors of HPPK have been reported to date. Early work, predating the structural characterization of HPPK, led to the identification of the gem dimethyl- (1) and phenethyl-substituted (2) substrate analogues as inhibitors (Figure 1).²⁹ Rationally designed bisubstrate analogues based on these have also been reported.^{30–32} Recently, adopting a rapid overlay of chemical similarity (ROCS) scaffold hopping screening method, we measured the binding of a series of commercially available substrate-like compounds and identified 8-mercaptoguanine (8MG) (Figure 1B) as a novel inhibitor of the *S. aureus* enzyme (SaHPPK) ($K_D = 11 \mu\text{M}$, $\text{IC}_{50} = 41 \mu\text{M}$).²¹ Unlike other inhibitors of HPPK, binding of this compound was shown to be independent of cofactor and Mg^{2+} ions. 8MG benefits from a high level of steric and electronic complementarity to the substrate site and being small, resulting in a high ligand efficiency

rating ($\Delta G/\text{number of heavy atoms} = 0.63 \text{ kcal mol}^{-1} \text{ heavy atom}^{-1}$). On the basis of this discovery, we initiated an ongoing SAR program to structurally elaborate 8MG into a more potent inhibitor.²⁰ Because of the demonstrated importance of the sulfur substituent, we initially focused on chemical extension from the N9 and N7 positions, revealing that only the latter strategy was viable. An N7-ethyl alcohol variant exhibited comparable affinity to that of the parent compound, but it was competitive with the cofactor.

In view of the above findings, we decided to revisit extension from the sulfur atom of 8MG. This article reports the results of this latest work, which has culminated in the identification of several novel 8MG analogues displaying significantly higher affinities ($K_D \sim 0.45 \mu\text{M}$) for SaHPPK compared to that of the parent compound in the presence of saturating levels of cofactor. The SaHPPK binding ability and inhibitory activity of an initial lead compound, as well as a series of synthesized analogues, have been quantified by surface plasmon resonance (SPR) experiments and an in vitro luminescent kinase coupled-enzyme assay, respectively, and important structural and dynamic features of the lead/cofactor/SaHPPK ternary complex have been revealed by heteronuclear NMR spectroscopy. By solving the X-ray structure of this complex, the intermolecular interactions between the lead inhibitor and the enzyme have also been delineated. Additionally, the X-ray structure of the wild-type SaHPPK with a cofactor analogue bound has been determined, revealing key residues required for cofactor recognition and those that have moved by over 30 Å upon binding of the inhibitor. Finally, SPR data is presented that show that a number of the new 8MG analogues are also able to bind to DHPS with low micromolar affinity, indicating the potential of this class of compound to be developed into dual-action enzyme inhibitors.³³

RESULTS AND DISCUSSION

A simple similarity search based on the TimTec (www.timtec.net) catalogue was performed, and the four commercially available 8MG analogues (out of seven hits) were purchased: 8-N-morpholinoguanine, 7-methylguanine, 8-bromoguanine, and 8-((2-(4-methoxyphenyl)-2-oxoethyl)thio)guanine (3). The SaHPPK binding properties of the first three of these compounds were reported in our earlier work.²⁰ Compound 3 is reminiscent

of the known 7-methyl-7-phenethylpterin analogue (2) of Wood et al.²⁹ (Figure 1B), which was crystallized, along with the 7-gemdimethyl variant (1), in the first EcHPPK structure by Stammers et al.³⁴

Binding of 3 to SaHPPK by SPR. Binding of 3 to SaHPPK was initially quantitatively analyzed by SPR, with SaHPPK immobilized on a nitrilotriacetic acid (NTA) sensor chip surface. Binding sensorgrams (Figure 2) were of good quality and

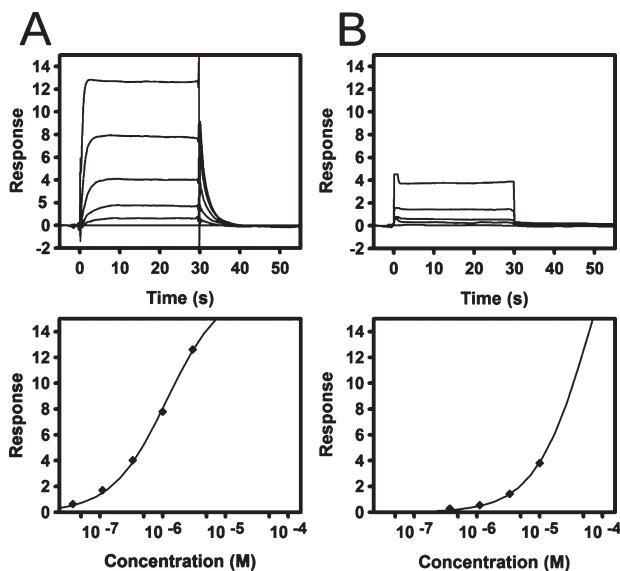


Figure 2. SPR sensorgrams (top panels) and steady-state affinity fits (bottom panels) for the binding of compound 3 to SaHPPK in the (A) presence or (B) absence of 1 mM ATP.

consistent with 1:1 stoichiometric binding. The binding affinities (K_D values) were derived by globally fitting the steady-state data sets to a single-site binding model. Thus, in the presence of ATP, 3 was estimated to bind SaHPPK with an affinity of $1.09 \pm 0.12 \mu\text{M}$. On the other hand, in the absence of ATP, steady-state fitting using 8MG as a reference ($K_D = 10.8 \pm 0.4 \mu\text{M}$)²¹ revealed far weaker binding ($K_D = 77 \pm 16 \mu\text{M}$).

Binding of 3 to SaHPPK by NMR Spectroscopy. The chemical shift of a nucleus is highly sensitive to changes in its local environment and is thus a convenient site-specific probe for analyzing ligand-binding events, including conformational and dynamic changes during complex formation.³⁵ Incremental addition of 3 to a sample of apo SaHPPK revealed only moderate strength binding, as evidenced by signal broadening (characteristic of intermediate exchange) and very minor chemical shift perturbations (CSPs) (data not shown). Broadened resonances mapped approximately to the substrate-binding site (data not shown). In stark contrast, titration of 3 into a solution of SaHPPK with either ATP or the nonhydrolyzable ATP analogue, AMPCPP, fully bound resulted in widespread chemical shift changes (Figure 3A), and all CSPs exhibited slow exchange on the NMR time scale. Moreover, in contrast to the case with 8MG,²¹ saturation was achieved at close to a 1:1 ligand-to-enzyme ratio, consistent with the higher affinity of 3 for SaHPPK measured by SPR (vide supra) in the presence of the nucleotide.

To investigate the binding of 3 in more detail, we assigned the ¹HN, ¹⁵N, and ¹³CA backbone resonances for the 3/AMPCPP/SaHPPK ternary complex using a triple-resonance 3D NMR experiment. The absolute CSPs induced by binding of 3 to the

AMPCPP bound SaHPPK were distributed mainly over two regions that, when mapped to the structure (Figure 3D), are consistent with pterin-site binding. Amides 8–12 on the sheet lining the pterin pocket showed small absolute CSPs around the one-standard deviation (σ) value (Figure 3B). Amides from residues 44–53 in loop 2 that are more solvent-exposed (see green or surface representation) and the more buried ring stacking amide from Phe123 showed even larger CSPs, over 3σ in magnitude. Other large CSPs were also observed for Gly90, Cys80, His82, and the side-chain H ϵ 2 of His115 in the loop 3 hinge region, which probably reflect changes in loop structure between the AMPCPP/SaHPPK binary and 3/AMPCPP/SaHPPK ternary complexes. Although the structure of the 8MG/AMPCPP/SaHPPK complex has yet to be determined, a similarly large CSP is observed for Gly90 upon the addition of 8MG (Figure 3A), suggesting that the position and environment of the tip of loop 3 in this structure is similar to that in the 3/AMPCPP/SaHPPK complex, at least around the Gly90 position.

In order to specifically probe the binding mode of the S8-substituted pendant in 3, together with any associated conformational changes in the SaHPPK structure, the CSPs for the 3/AMPCPP/SaHPPK complex were compared to those derived from the 8MG/AMPCPP/SaHPPK spectra (Figure 3C) and depicted in a model constructed from the AMPCPP/SaHPPK X-ray structure (vide infra). As part of this initial analysis, it was assumed that the guanine moieties of 8MG and 3 superpose in the substrate pocket, and the loop 2 conformation was derived from that observed in the 8MG/SaHPPK binary complex.²¹ Differential CSP data is quite powerful for determining the pose of bound ligands by NMR,³⁶ and this analysis showed that the pendant of 3 interacts closely with residues in loop 2, specifically around Tyr48, extending up and out from the base of the substrate-binding site (Figure 3D). The differential CSP observed for Arg121 may be suggestive of a direct interaction with the pendant. The differential CSPs that mapped to those residues under the pterin pocket (amides 8–10), remote from the pendant, may support a small change in the binding orientation of the guanine moiety compared to that in the 8MG/AMPCPP/SaHPPK case. Although Asp95 and His115 are not solvent-exposed, the changes in the CSPs reflect a change in the environment near the gamma phosphate, potentially due to loop 3 structural changes (vide infra).

X-ray Structures of 3/AMPCPP/SaHPPK and AMPCPP/SaHPPK. The X-ray structure of SaHPPK in complex with 3 and AMPCPP (Figure 4) was solved at 2.0 Å resolution using molecular replacement (Table 1). The ternary complex crystallized in the $P6_1$ space group, with a single protein molecule in the asymmetric unit. Backbone density was observed for all 158 amino acid residues of the protein. The guanine moiety of 3 is positioned similarly to that of 8MG in the 8MG/SaHPPK binary complex,²¹ making a total of six hydrogen bonds with the protein and π -stacking between the aromatic rings of Phe54 and Phe123. The aromatic ring of the pendant projects out and away from the substrate pocket into the loop 2/loop 3 region, making favorable hydrophobic interactions with Val46 and Gly47. The adjacent ketone group stacks against the guanidinium group of Arg121 and also interacts with the phenyl ring of Phe123. These features are also consistent with the large CSPs observed for these residues upon formation of the ternary complex (vide supra).

The Mg²⁺ ions in the 3/AMPCPP/SaHPPK structure superpose closely with those in the HMDP/AMPCPP/EcHPPK structure (PDB: 1QON),^{22,23} differing in position by only 0.22 (Mg1) and 0.35 Å (Mg2). Both ions are coordinated to residues

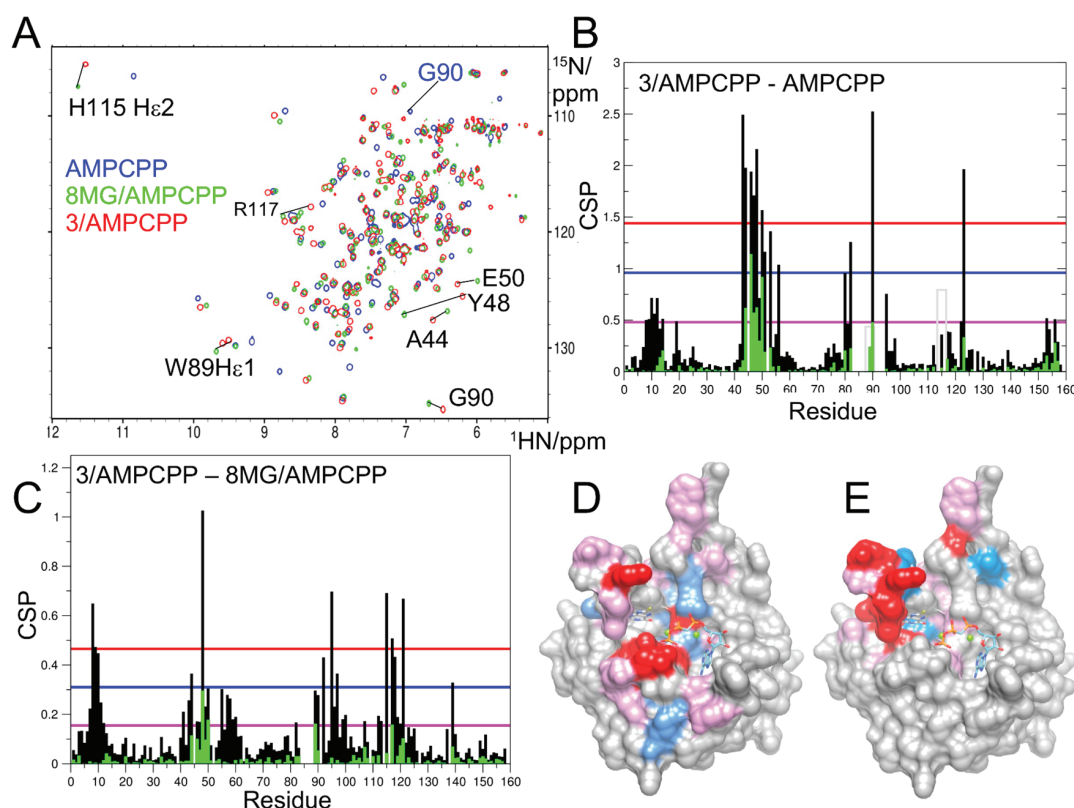


Figure 3. NMR data of SaHPPK binding to 3. (A) Superposition of the 2D ^{15}N HSQC spectra recorded in a sample of $\sim 120\ \mu\text{M}$ SaHPPK in 10 mM Mg^{2+} 50 mM HEPES, pH 8, in the presence of saturating amounts of AMPCPP (blue), 8MG (green), and compound 3 (red). (B) Raw CSP (black) for the change in weighted averaged chemical shifts for 3/AMPCPP compared to the AMPCPP 2D ^{15}N HSQC spectra. CSP plot weighted by solvent accessibility (green). CSPs for the side chain of His115 and Trp89 are shown in gray. (C) Raw weighted average chemical shifts (black) derived from the change in the 3/AMPCPP and the 8MG/AMPCPP 2D NMR spectra and weighted for solvent accessibility (green). In both panels B and C, pink, blue, and red horizontal lines signify 1, 2, and 3 standard deviations of the CSPs. (D, E) CSPs greater than the values in panels B and C are mapped to the structure of SaHPPK in panels D and E, respectively.

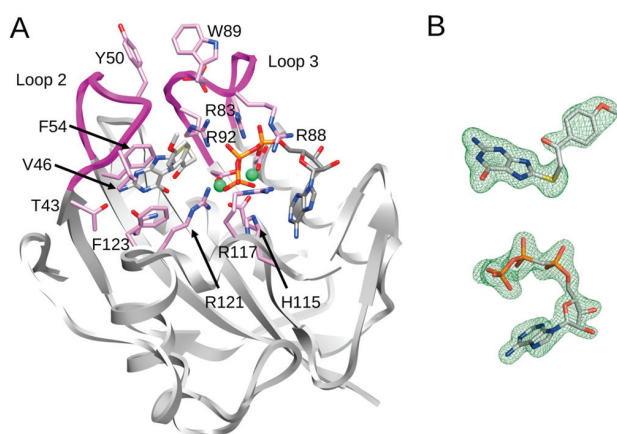


Figure 4. X-ray structure of SaHPPK in complex with AMPCPP and 3. (A) Detail of the active site. Loops 2 and 3 are shown in magenta, and the two magnesium ions, in green. (B) $mF_o - DF_c$ difference density map of AMPCPP and 8MG contoured at 3.0σ .

D95 and D97 and share coordination to a β -phosphate oxygen. One magnesium center is coordinated by the α -phosphate, whereas the other is bound by the γ -phosphate. The fifth and sixth coordination sites are occupied by water molecules in each case.

To help rationalize the conformational changes accompanying binding of 3, as well as the ~ 70 -fold enhanced affinity of 3 for the cofactor/SaHPPK complex relative to that for SaHPPK alone, the X-ray structure of SaHPPK in complex with AMPCPP was also determined (Figure 5). The complex crystallized in the space group $P2_1$ (Table 1) and was solved via molecular replacement to a resolution of $2.7\ \text{\AA}$. In contrast to the 8MG/SaHPPK complex, four rather than two protein molecules were found in the asymmetric unit (Figure 5A).²¹ There are also differences between the two crystal forms in terms of the nature of the protein–protein interface. The existence of an intermolecular disulfide bond between the solvent-exposed Cys80 residues of neighboring proteins in the AMPCPP/SaHPPK crystal lattice is particularly notable (Figure 5B). Despite the presence of AMPCPP in the crystallizing solution used to grow crystals of the 8MG/SaHPPK binary complex, AMPCPP did not bind in this case, which was rationalized in terms of the binding site being partly occluded by the interface of the protomers in the asymmetric unit. The change in the nature of the protein–protein interface associated with intermolecular disulfide bond formation appears to be more compatible with AMPCPP binding.

As anticipated, the AMPCPP occupies the cofactor site, with phosphate oxygens hydrogen-bonded to Arg121, Arg117, His115, and Arg92 (one of three arginines in loop 3). The adenine base is hydrogen-bonded to the amide backbone of Ile98

Table 1. X-ray Statistics

	AMPCPP (4CYU)	AMPCPP/3 (4CRJ)	AMPCPP/7 (4CWB)
spacegroup	$P2_1$	$P6_1$	$P6_1$
wavelength (Å)	1.0080	0.9537	0.9537
unit-cell parameters (Å, deg)	$a = 62.00, b = 94.26, c = 62.99, \alpha = \gamma = 90.00, \beta = 112.44$	$a = b = 82.48, c = 52.17, \alpha = \beta = 90.00, \gamma = 120.00$	$a = b = 83.82, c = 52.04, \alpha = \beta = 90.00, \gamma = 120.00$
Diffraction Data			
resolution range (Å)	47.13–2.70 (2.77–2.70)	42.17–2.00 (2.05–2.00)	42.33–1.56 (1.60–1.56)
no. of unique reflections	18 506 (2455)	13 792 (1962)	29 793 (1419)
no. of observed reflections	139 533	296 051	332 946
Matthews coefficient, V_M (Å ³ Da ^{−1})	2.39	2.83	2.93
solvent content (%)	48.6	58.6	58.1
completeness (%)	100.0 (100.0)	99.7 (98.0)	99.7 (96.9)
data redundancy	7.5 (7.6)	21.5 (19.7)	11.2 (10.3)
mean $I/\sigma(I)$	9.4 (2.1)	12.2 (3.6)	14.0 (2.7)
R_{merge}	0.139 (0.899)	0.215 (1.02)	0.091 (0.673)
$R_{\text{p.i.m.}}$	0.082 (0.528)	0.047 (0.232)	0.041 (0.315)
Refinement			
R_{free} (%)	24.2	20.0	17.8
R_{cryst} (%)	19.4	16.4	16.0
size of R_{free} set (%)	4.5	5.0	5.1
protein molecules in the asymmetric unit	4	1	1
inhibitor molecules		1	1
cofactor-analogue molecules	4	1	1
water molecules	13	83	141
RMSD from Ideal Values			
bond lengths (Å)	0.010	0.006	0.009
bond angles (deg)	1.49	1.36	1.60
Mean B factors (Å ²)	54.2	27.4	15.5
Ramachandran Plot			
favoured (%)	99.0	98.7	98.2
outliers (%)	0.2	0.0	0.0

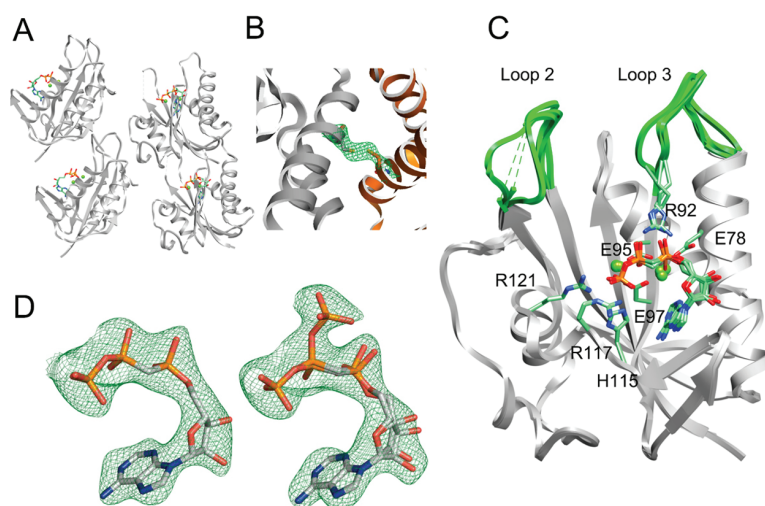


Figure 5. X-ray structure of SaHPPK in complex with AMPCPP. (A) X-ray structure showing the spatial arrangement of the four protomers in the asymmetric unit. (B) $2mF_o - DF_c$ standard density map (green mesh) contoured at 1.5σ showing a disulfide bond observed between protomers in adjacent unit cells. (C) $mF_o - DF_c$ difference density map (green mesh) of AMPCPP from protomers A (left) and B (right) contoured at 3.0σ . (D) Superposition of the four protomers and detail of the interactions of the bound AMPCPP. In panels A and C, alternate conformations of AMPCPP are omitted for clarity, and magnesium ions are colored green.

and Ser112 and forms hydrophobic contacts with the side chains of Leu111, Ile98, and Leu71. The ribose O2' is hydrogen-bonded to Lys110. Two magnesium ions are again present, coordinated by Asp95, Asp97, and either the α - and β -phosphates or the β -

and γ -phosphates of AMPCPP. Alternate conformations of AMPCPP are present in protomers B and D. In these conformations, the γ -phosphate is oriented orthogonal to the γ -phosphate site (Figure 5D), forming additional interactions to

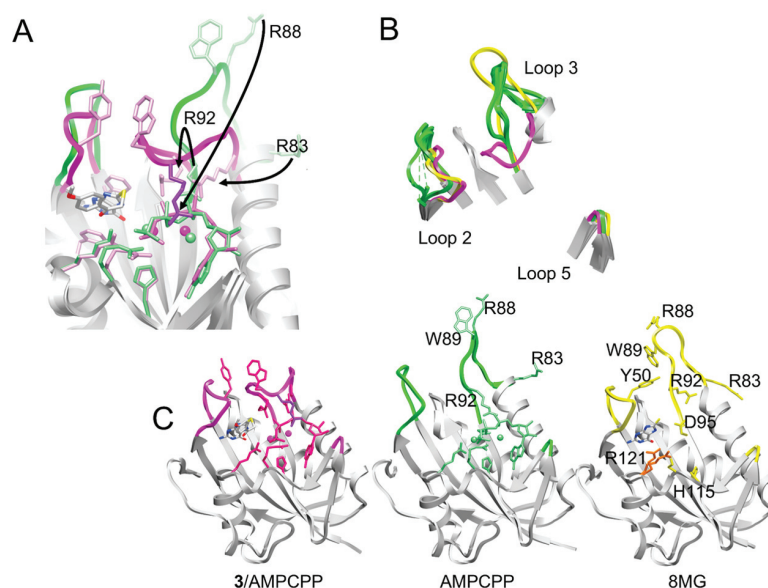


Figure 6. Conformational changes in SaHPPK. (A) Superposition of the AMPCPP (green) and the AMPCPP/3 (magenta) X-ray structures illustrating the change in conformation of the loop 3 and selected arginine side chains. (B) Superposition of the three SaHPPK X-ray structures shown in panel C highlighting the changes in the loop regions. (C) Comparison of the X-ray structure of SaHPPK in complex with AMPCPP/3, AMPCPP, and 8MG. Selected side chains are shown to illustrate the positional and conformational changes.

Arg92 and Gln3 of an adjacent protomer, at the expense of loss of hydrogen bonds to His115, Arg118, and Arg121 as well as coordination to a magnesium ion.

Density was either weak or not observed for several of the loop 2 residues, indicative of mobility therein, concordant with prior NMR relaxation studies on EcHPPK³⁷ and SaHPPK²¹ in which loop 2 was shown to be dynamic in the absence of bound substrate or substrate inhibitor. The observed loop 3 conformations (Figure 5C) overlap well with those previously found in the ensemble of NMR structures of AMPCPP/EcHPPK (data not shown),²⁴ and the existence of multiple conformations in the structure is again in line with the broadened NMR resonances observed for residues 84–89 and 92 (vide supra). In the AMPCPP complex, the side chain of Arg92 is extended and makes an end-on salt bridge to the α - and β -phosphates, which caused loop 3 to adopt a more open conformation relative to that observed in the crystal structure of apo EcHPPK.

It is noted that the side-chain orientation of Arg92 is not at all defined in the NMR structure of EcHPPK, which has been used in the past to represent the EcHPPK/AMPCPP structure in the absence of a suitable EcHPPK X-ray structure,²³ although this may reflect, to some degree, the difficulty and limitation of employing short-range NOEs to characterize hydrophilic interactions. The important role of Arg92 in cofactor recognition is underscored by the current AMPCPP/SaHPPK complex, which resembles the predicted AMPCPP/EcHPPK model derived from locally enhanced sampling and molecular dynamics simulations.³⁸

A comparison of the two structures presented here with the previously reported 8MG/SaHPPK structure (PDB: 3QBC) (Figures 6A–C)²¹ reveals marked differences in the conformation of the cofactor loop 3 and the substrate loop 2 and more subtle change in loop 5 (Figure 6B). There is a notable change in position of Arg92 upon binding of 3. In the 8MG/SaHPPK structure, the guanidinium group of Arg92 is displaced by that of Arg83 and is translated 3.5 Å toward the substrate pocket. Arg83 moves 13 Å to make a hydrogen-bond contact to the α -

phosphate. In the 3/AMPCPP/SaHPPK complex, the Arg92 headgroup is also rotated by 90° and is oriented approximately parallel to that of Arg83; it hydrogen bonds to the β -phosphate as before. Gly90 at the tip of loop 3 moves ca. 17 Å closer to the active site upon the binding of 3 to the AMPCPP/SaHPPK complex. The conformations of the side chains of the other residues involved in nucleotide binding (Leu71, Leu75, Glu78, Asp95, Asp97, Ile98, Lys110, Leu111, Ser112, Val113, His115, and Arg117) do not change between the two structures.

The AMPCPP lies over 4 Å away from 3 in the ternary complex; therefore, the significantly enhanced binding affinity of 3 toward SaHPPK in the presence of the cofactor must be due to cofactor-induced intermolecular interactions rather than specific interactions with the cofactor itself. In the 8MG/SaHPPK structure,²¹ residue Arg121 exists in two conformations (shown in orange in Figure 6C), one oriented toward and one perpendicular to the cofactor, suggesting that this residue is likely mobile in the apo enzyme. In both the AMPCPP/SaHPPK and 3/AMPCPP/SaHPPK structures, the conformation appears to be locked by the hydrogen bond to the γ -phosphate of the AMPCPP, possibly assisted by a π interaction between the guanidinium group of Arg121 and the ketone of 3, resulting in the observed cofactor-mediated improvement in binding affinity.

From the 3/AMPCPP/SaHPPK structure, the large, yet similar, chemical shift change observed in the NMR spectra for Gly90 upon binding of 3 (Figure 3A) or 8MG to AMPCPP/SaHPPK can be attributed to hydrogen bonding to the carbonyl of Glu87 combined with a favorable interaction between the sulfur atom of each ligand and the carbonyl group of Trp89,³⁹ rather than a previously hypothesized hydrogen bond between the SH of 8MG and the carbonyl of Trp89.²¹

Insight into the Dynamics of the 3/AMPCPP/SaHPPK Complex by NMR. To investigate the dynamic properties of the 3/AMPCPP/SaHPPK complex on the fast (pico- to nano-second) time scale, ¹⁵N heteronuclear NOEs were measured, and those amides with NOE values chosen to be less than 0.75 were mapped onto the surface of the structure (Figure 7) to highlight

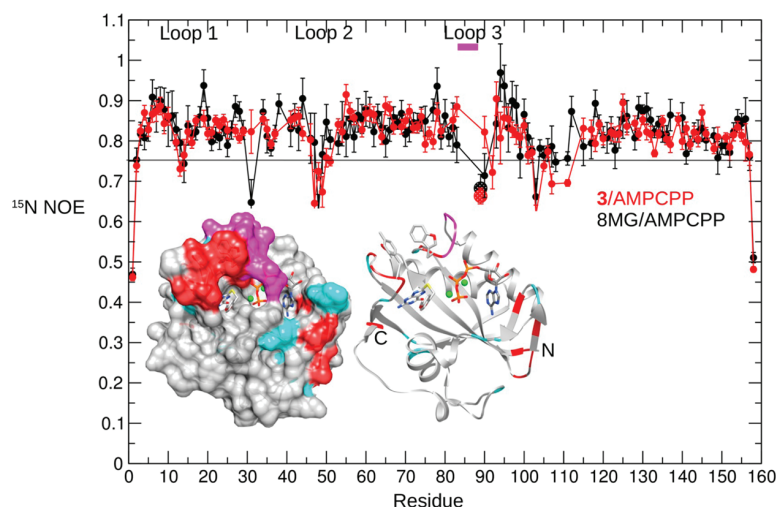


Figure 7. 600 MHz ^{15}N heteronuclear NOE values for SaHPPK in complex with AMPCPP/3 (red) and AMPCPP/8MG (black). The ^{15}N NOE for the side-chain H ϵ 1 for Trp89 is shown as a large shaded circle. Residues with ^{15}N NOE values less than 0.75 are mapped onto the surface and ribbon representation of the SaHPPK/AMPCPP/3 X-ray structure. Amides not observed in the ^{15}N HSQC spectra due to severe broadening in loop 3 are colored magenta, and proline residues, cyan.

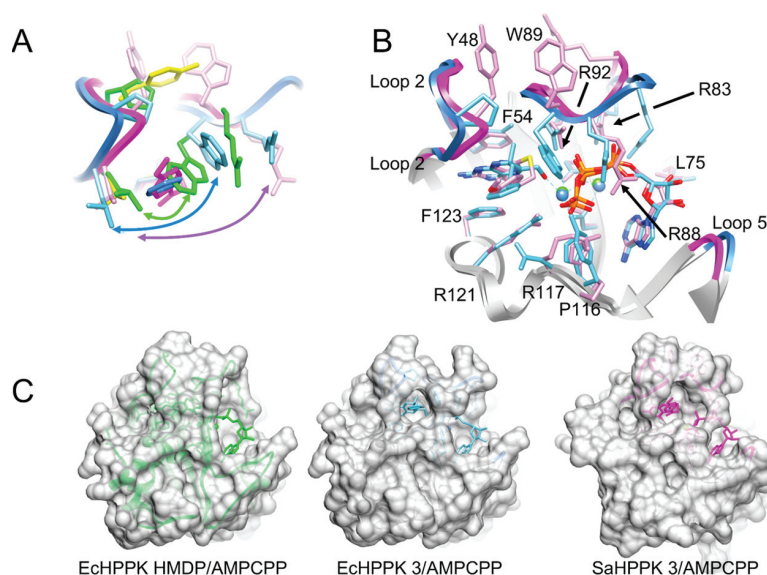


Figure 8. Comparison of the SaHPPK AMPCPP/3 structure with the EcHPPK/AMPCPP/2 structure. (A) The substrate pocket in the SaHPPK/AMPCPP/3 complex (magenta) is more open than that in EcHPPK/AMPCPP/2 (blue) or in EcHPPK/AMPCPP/HMDP (green), which is completely sealed. (B) Comparison of the active site of SaHPPK/AMPCPP/3 (magenta) with that of EcHPPK/AMPCPP/2 (blue). (C) Surface representation of the EcHPPK/AMPCPP/HMDP (left), EcHPPK/AMPCPP/2 (middle), and SaHPPK/AMPCPP/3 complexes (right).

statistically relevant regions with increased mobility compared to the global average (0.81). Residues 1 and 158 at the termini are highly mobile (^{15}N NOE < 0.5). A characteristic dip in the ^{15}N NOE values from the global average for amides 47–51 is indicative of some residual fast-time scale motion centered around the interface of loops 2 and 3. Other amides experiencing limited motion include Arg92 in the hinge of loop 3 and Glu103, located in a loop leading into the β hairpin, within which Asp107 and Leu111 are also partly mobile.

The data were compared with that for the 8MG/AMPCPP/SaHPPK complex (shown in black) by recording the sample under identical conditions and with the same pulse sequence. A combination of a newer cryoprobe and a more optimal pulse sequence afforded much better water suppression than

previously achieved for 8MG/AMPCPP/SaHPPK²¹ and yielded very similar results for both ternary complexes, with the standard deviation of the average value noticeably smaller than that reported earlier (0.08 vs 0.12). The ^{15}N NOE for the side chain of Trp89 appears to be essentially the same for both the 3/AMPCPP/SaHPPK and 8MG/AMPCPP/SaHPPK complexes, with a degree of residual fast motion similar to that observed for loop 2, indicating that the interaction of the pendant of 3 with this residue is not strong enough to dampen Trp89's fast time scale motion entirely. From a structural perspective, this observation may be more consistent with the position of the Trp89 side chain observed in the X-ray structure compared to that observed in the HMDP/AMPCPP/EcHPPK ternary complex.⁴⁰

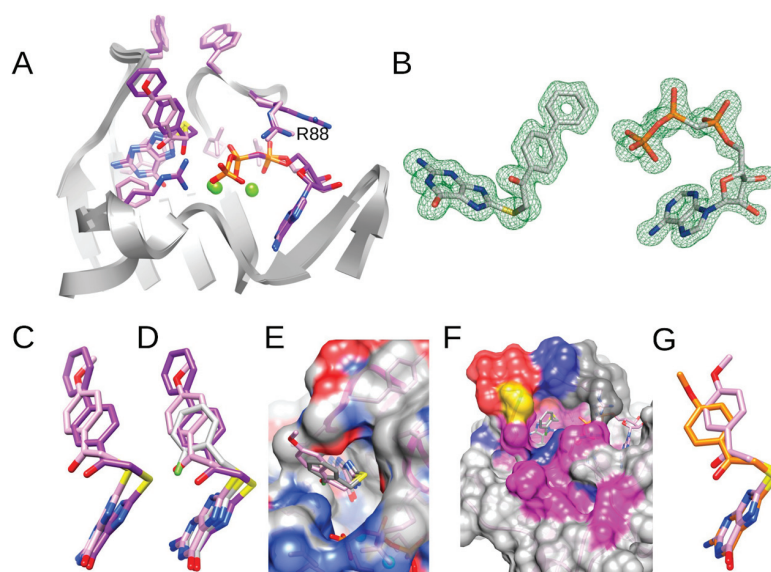


Figure 9. (A) Overview of the active sites of the SaHPPK/AMPCPP/3 (pink) with SaHPPK/AMPCPP/7 complexes and (B) the $mF_o - DF_c$ difference density map of the ligands in the SaHPPK/AMPCPP/7 complex contoured at 3.0σ . (C) Superposition of the bound poses for 3 (pink) and 7 (magenta) with (D) the modeled bound pose of 11 (white). (E) 3 and 11 shown within a surface representation of the substrate pocket. (F) Raw CSPs corresponding to the change in the weighted average chemical shifts observed for 11/AMPCPP/SaHPPK relative to 8MG/AMPCPP/SaHPPK in the 2D ^{15}N HSQC NMR spectra, mapped onto the docked 11/AMPCPP/SaHPPK structure. The CSPs are colored pink, blue, and red for those greater than the 1, 2, or 3 times the standard deviation of the CSP values shown in Figure 3C. The broadened Gly47 is colored yellow. (G) The bound pose of 3 in the EcHPPK/AMPCPP/3 (PDB: 4M5J, orange) is different from that in the SaHPPK/AMPCPP/3 structure (pink).

Even though 3 is larger than 8MG, binds more tightly to the AMPCPP/SaHPPK complex (by over an order of magnitude), and protrudes into the loop 2/3 region, amide signals for residues 84–89 in loop 3 were not observed in the NMR spectra of 3/AMPCPP/SaHPPK (magenta in Figure 7), revealing large amplitude motion on the slower (micro- to millisecond) time scale. Thus, the addition of the pendant to the 8MG parent scaffold appears to have limited large-scale impact on the overall micro- to millisecond backbone dynamics of loop 3. However, given that the pendant forms few intermolecular interactions with side-chain atoms, this is not entirely unexpected.

Comparison of the Structures of 3/AMPCPP/SaHPPK, 2/AMPCPP/EcHPPK, and 3/AMPCPP/EcHPPK. Given the broad similarity in the structures of inhibitors 2 and 3 and the sequence similarity between EcHPPK and SaHPPK, it is not surprising that the 3/AMPCPP/SaHPPK structure is generally quite similar to that of the previously reported 2/AMPCPP/EcHPPK ternary complex (PDB: 1DY3) (Figure 8A,B),³⁴ with the cofactor loop 3 closed in over the active site in both cases, compared with the extended conformation observed in the 8MG/SaHPPK binary complex.²¹

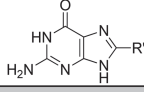
In the case of the *E. coli* enzyme, a comparison of the 2/AMPCPP/EcHPPK complex with the substrate/AMPCPP/EcHPPK complex⁴⁰ shows that the binding of 2 leads to an opening of the substrate pocket lid, revealing a solvent-exposed active site pocket, filled by the phenethyl pendant of 2. This widening to accommodate 2 causes a rotation and 3.6 Å positional change in the side-chain methyl groups of Leu45 and a 3.1 Å sized hinge movement of Trp89 (Figure 8A). Although the equivalent SaHPPK structure in complex with HMDP and AMPCPP has not been crystallized, a comparison with the 8MG structure²¹ reveals that the rotation of the equivalent Val46 methyl groups is significantly smaller (<1 Å) in comparison (Figure 8A).

Although the phenyl rings of 2 and 3 in the enzyme complexes occupy similar positions, the position of the side chain of the Trp89 residue is markedly different in the two structures (Figure 8B). In the 2/AMPCPP/EcHPPK structure, the Trp89 indole ring orients toward the binding site, forming an edge-on π -stacking interaction with the phenyl ring of 2 on one side and the Arg88 residue on the other, and the indole ring He2 atom makes a hydrogen-bond contact to the γ -phosphate of ATP. In the SaHPPK structure, however, Trp89 orients away from the binding site to form a hydrophobic contact with the loop 2 residue, Tyr48 (Pro in EcHPPK). The Arg88 guanidinium group is translated 5 Å deeper into the cofactor site, where it appears to hydrogen bond to the heterocyclic ribose oxygen of AMPCPP.

For EcHPPK, formation of the HMDP/AMPCPP ternary complex involves a hydrogen-bond network that depends on the interaction of residues Asn11 and Gln51 (equivalent to Asn10 and Gln50 in SaHPPK) to draw all three loop regions into the fleeting transition state conformation.²³ While we have not yet solved the structure of the HMDP/AMPCPP/SaHPPK complex, it is possible that a Trp89-Tyr48 interaction helps to stabilize the interaction between loops 2 and 3 within this complex. A surface representation of all three ternary complexes (Figure 8C) illustrates that the positioning of the Trp89 and Arg88 residues in the 3/AMPCPP/SaHPPK complex leads to a larger, more solvent-accessible binding pocket than that in the 2/AMPCPP/EcHPPK complex. In contrast, the substrate-binding site is completely shielded from solvent in the HMDP/AMPCPP/EcHPPK complex.⁴⁰

Comparing the cofactor sites, in EcHPPK, three arginines (Arg82, Arg84, and Arg92) bind the phosphates of AMPCPP in 2/AMPCPP/EcHPPK. For SaHPPK, Arg83 and Arg92 replace the roles of Arg84 and Arg92 in EcHPPK; however, Arg85 points in the opposite direction to the equivalent Arg84 in EcHPPK (Figure 8B). This difference may be due to the single-residue

Table 2. SAR for Compounds 3–13

				
Compound	R'	SaHPPK ^a <i>K_D</i> (μM)	SaHPPK ^b <i>K_D</i> (μM)	EcDHPS ^b <i>K_D</i> (μM)
8MG	SH	12.6 ± 1.3	10.8 ± 0.4 ^c	76.2 ± 12
3		1.1 ± 0.1	76.6 ± 17	3.9 ± 0.2
4		1.4 ± 0.2	10.3 ± 3.6	3.9 ± 0.7
5		1.9 ± 0.7	7.5 ± 0.1	8.2 ± 2.4
6		2.8 ± 0.4	21.7 ± 3.8	1.6 ± 0.5
7		0.81 ± 0.01	21.5 ± 1.1	8.4 ± 1.3
8		7.8 ± 0.50	73.5 ± 6.40	14.2 ± 0.40
9		1.85 ± 0.07	42.7 ± 11.2	7.8 ± 1.0
10		1.79 ± 0.69	13.5 ± 0.21	43.9 ± 9.8
11		0.45 ± 0.095	4.3 ± 0.6	8.5 ± 2.1
12		2.81 ± 0.14	7.26 ± 0.12	9.74 ± 0.88
13		0.66 ± 0.035	11.1 ± 1.3	12.3 ± 0.6

^aWith 1 mM ATP present. ^bNo ATP present. ^cChhabra et al.²¹

insertion in loop L3 in the *E. coli* enzyme (Ala86). Arg84 in EcHPPK hydrogen bonds to the α -phosphate of AMPCPP, resulting in a difference of 1.2 Å in the positioning of the O5' atoms in the two ternary structures. The position of the O3' atoms also differs by 2.3 Å, likely due to interaction with Gln74 in EcHPPK and Leu75 in SaHPPK, which alters the ring pucker (C2' endo in SaHPPK and C3' endo in EcHPPK). The change in the O2' positions in the two puckers and the hydrogen bond from O2' to the backbone of Lys110 in SaHPPK (Arg110 in EcHPPK) may explain the observed 2.5 Å difference in the positioning of the tip of the cofactor loop 5 (Figure 8B).

During the later stages of this work, Yun et al.⁴¹ also reported the X-ray structure of 3 in complex with AMPCPP/EcHPPK. A comparison of the 3/AMPCPP/SaHPPK and 3/AMPCPP/

EcHPPK structures reveals some interesting differences pertaining to the active site region. Most notably, the active site is more solvent-exposed in the SaHPPK structure, and the pendant of 3 adopts different poses, with the ketone oxygen and para carbon of the pendant ring differing in position by 1.6 and 1.2 Å, respectively, between the two structures (Figure 9F). Given that the NMR data for the 3/AMPCPP/SaHPPK complex indicates that both residues are in close contact with the ligand and that the loops surrounding the pendant of 3 are not completely rigidified on either the fast or slow time scale, it would be interesting to conduct a similar study on 3/AMPCPP/EcHPPK.

SAR of Compound 3 and Other S8-Substituted Guanine Analogues. To probe the contribution of specific

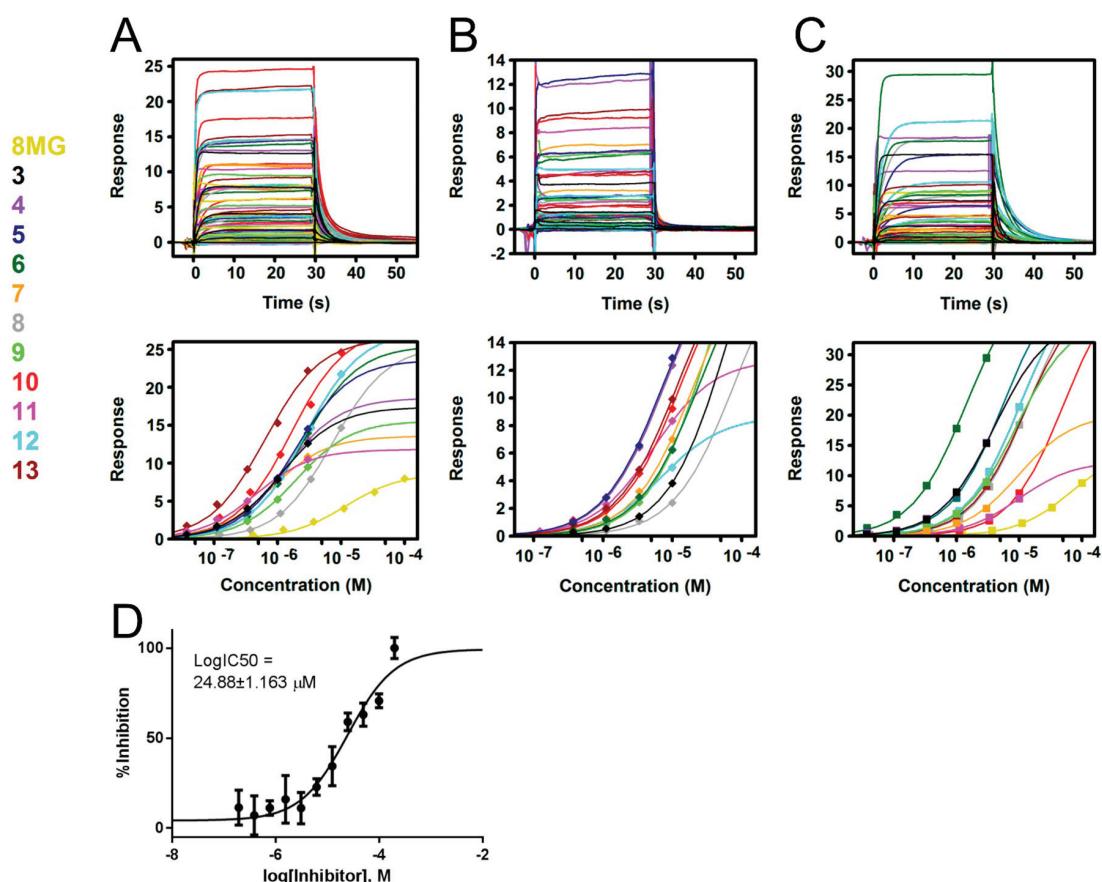


Figure 10. SPR sensorgrams (top panels) and steady-state affinity fits (middle panels) for the binding of all compounds to SaHPPK in the (A) presence or (B) absence of 1 mM ATP and to (C) EcDHPS. (D) Kinase-Glo assay results for compound 11. Error bars indicate SEM.

groups to the binding affinity of 3 toward SaHPPK, we embarked on a SAR investigation and synthesized a series of S8-substituted 8MG analogues. The affinity and activities for the series of compounds are shown in Table 2 and Figure 10.

We first investigated the likely contributions of the methylene, aryl, carbonyl and methoxy groups of 3 ($K_D = 1.1 \mu\text{M}$) toward its affinity for the cofactor-bound enzyme. Removal of the 4-methoxy (compound 4, $K_D = 1.4 \mu\text{M}$) had little effect on binding, an observation predicted from the solvent-exposed nature of this group in the X-ray structure (Figure 3A). A branched methylene was also tolerated (5), furnishing a $K_D = 1.8 \mu\text{M}$, whereas moving the 4-methoxy group to the 2-position (6) gave a slight loss of affinity ($K_D = 2.8 \mu\text{M}$). Furthermore, replacement of the 4-methoxy group with a phenyl group to give biphenyl 7 did not significantly alter affinity ($K_D = 0.81 \mu\text{M}$). Saturation of the ketone group to give the methylene analogue 8, on the other hand, led to a significant decrease in affinity ($K_D = 7.8 \mu\text{M}$), which is again supported by the structural data, which revealed a specific interaction between the ketone group and the side chain of Arg121 in 3. Replacement of the aryl ring with a methyl group (9) led to only a slight decrease in affinity ($K_D = 1.9 \mu\text{M}$), suggesting that the interactions of the aryl ring with the enzyme are not strong, consistent with the relatively open or dynamic nature of the active site around the aryl group.

In their recent study with EcHPPK, Yun et al.⁴¹ also found that the 4-OMe group in 3 was not important for binding to EcHPPK. However, in contrast to our findings for SaHPPK, they found that a branched methylene linker and moving the 4-methoxy to

the 2-position (6) led to complete abrogation of activity, which is again suggestive of a more open active site around the aryl group in SaHPPK. Furthermore, they found that large groups at the 4-position were not tolerated, whereas we have shown that they can be for SaHPPK. Indeed, we were able to crystallize and solve the structure of the biphenyl analogue 7 in complex with AMPCPP (Figure 9A,B). The structure is remarkably similar to that of the 3/AMPCPP/SaHPPK complex. The largest differences are found in the orientation of the Arg88 side chain, which is split between two conformations. The difference of most relevance to the SAR analysis, however, is the change in the position of the ketone group (2.2 Å), which maintains the stacking between the Arg121 guanidinium group and the sulfur atom (Figure 8A,C). The position of the sulfur in the biphenyl analogue is essentially the same as that in the 8MG/SaHPPK binary structure.

In view of the observed changes in the sulfur and carbonyl positions (Figure 9C) in the cofactor-bound SaHPPK complexes of 3 and 7, it was reasoned that removal of a carbon from the linker in 3 might serve to draw the aryl ring down into the position observed in the 7/AMPCPP/SaHPPK structure. We reasoned that π stacking would be possible via the phenyl ring tethered to a shortened linker and that this could replace the ketone π system. Synthetic efforts were thus directed toward shortening the linker in 3 to a substituted benzyl group. Listed in Table 2 are the results for the unsubstituted compound 10, which effectively represents a direct replacement of the acetyl group of 9 with a phenyl ring. Gratifyingly, activity was maintained ($K_D = 1.8 \mu\text{M}$). A limited set of analogues was then assembled, the most

interesting of which are shown in Table 2. Here, it can be seen that a 2-CF₃ group results in a relative loss of affinity (**12**, K_D = 2.8 μ M). However, the 2-fluorobenzyl analogue **11** displayed an increase in affinity toward SaHPPK, with a K_D of 453 nM. Another compound of interest was the 4-cyano analogue **13**, which also displayed strong affinity toward SaHPPK with a K_D of 660 nM. Compared with **3**, compounds **11** and **13** benefit from a reduced molecular weight and number of rotatable bonds. For these reasons, in addition to the comparative ease of synthesis of these and other S-benzyl-substituted compounds, compounds such as **11** and **13** represent promising leads for further development. Interestingly, **11** also binds appreciably to the apo enzyme and does so to a far greater extent than that of **3**, with respective K_D values of 4.3 and 77 μ M. It may therefore be a good synthetic starting point for investigating cofactor-competitive binders or moieties that bind at the metal site, perhaps by combining SAR from our previously reported cofactor-competitive N7 ethyl alcohol 8MG analogue (K_D = 10 μ M for the apo enzyme).²⁰ Compound **13** also corresponds to the only S-benzyl-substituted analogue of 8MG tested by Yun et al.⁴¹ In contrast to our findings for SaHPPK, this compound exhibits very poor affinity for EcHPPK.

Crystallization attempts with **11** were unsuccessful; therefore, docking of **11** into the SaHPPK crystal structure was undertaken (Figure 9D, E). The resulting model suggests that the fluorine atom could position itself relative to the guanidinium group in a manner similar to that of the ketone group in **3**. This could plausibly confer increased affinity, as the guanidinium group is known to be highly fluorophilic and can interact strongly with a negatively polarized fluorine, favoring parallel orientations to the guanidinium plane,⁴² as observed in the docked model. The model indicates that the phenyl ring can overlay with the first ring in the biphenyl analogue **7**, maintaining interactions to Val46. A reason for the lower binding affinity of **12** could be due to the likely propensity for the 2-CF₃ group to twist the phenyl ring out of plane. The ¹⁵N HSQC spectrum of **11**/AMPCPP/SaHPPK was assigned using the HNCA experiment, as was done for the spectrum of **3**/AMPCPP/SaHPPK. Comparing the differential CSPs of 8MG/AMPCPP-saturated SaHPPK with either **3**/AMPCPP/SaHPPK (Figure 3C) or **11**/AMPCPP/SaHPPK (Figure 9F and Supporting Information Figure S1) reveals very similar pattern of CSPs but with notably larger differential CSPs for Tyr48 and Glu50 amides in loop 2. Of note, the signal for Gly47, the nearest amide to the pendant of **11** in the model, was not observed in the spectra (shown in yellow in Figure 9F) but was in the complex with **3**, most likely due to subtle motional effects of the proximal phenyl ring of the pendant on a micro- to millisecond time scale. The larger differential CSPs (>3 σ) of Tyr48 and Glu50 (shown in red) may represent a conformational change in loop 2, which also effects the stacking interaction with Trp89, as evidenced by differential CSPs of >2 σ (shown in blue). Substantial differences were also noted in the arginine side-chain region of the spectra (data not shown), which may be consistent with the Arg121 guanidinium–fluorine interaction predicted by the model. Although the observed NMR chemical shift data is supportive of the docked pose of **11**, assignment of the NMR signals of the side-chain atoms will be a prerequisite to determine the precise bound structure of **11**.

The binding data for **13**, **6**, and **7** highlights a clearly divergent SAR trend for SaHPPK and EcHPPK. The higher affinity of the latter two compounds for SaHPPK can be reconciled on the basis of the observed increased size/plasticity of the binding pocket for SaHPPK. In the case of **13**, the para-attached cyano group

polarizes the aryl ring, which would favor the π interaction with the Arg121, presumably for both enzymes. It is unclear why this would decrease the affinity compared to that of **3** for the EcHPPK enzyme. The enhanced affinity for SaHPPK may have its origins in the different pendant poses for **3** bound to the two enzymes (Figure 9G) as well as differences in the orientation of the methyl groups of Leu45 in EcHPPK and Val46 in SaHPPK (Figure 8A). In order to assess whether enhanced affinity translated into enhanced functional inhibition of HPPK, we selected **11** and tested it alongside 8MG (IC_{50} = 41 μ M)^{20,21} to obtain the IC_{50} value. This gave rise to a value of 25 μ M (Figure 10), suggesting that, while the higher affinity of **11** for HPPK gives rise to increased functional inhibition, the increase is not as great as might have been expected in view of **11**'s much higher affinity for HPPK (in the presence of ATP). However, it is noted that in terms of drug-likeness, our lead compound **11** has a number of favorable properties. For a compound with submicromolar affinity, it has a relatively low molecular weight of 291 Da, a topological polar surface area (tPSA) of 95 Å², suitable for membrane permeability and oral availability,⁴³ and a cLogP of 3.5. On the downside, it contains four hydrogen-bond donors, and, accordingly, its membrane penetration ability was found to be poor ($A-B P_{app}$ = $0.6 \pm 0.2 \times 10^{-6}$ cm s⁻¹ for Caco-2 monolayers).

Binding of Compounds to DHPS. Given the chemical similarity of the guanine and pterin scaffolds, the fact that the pterin core is common to both HPPK and DHPS substrates, and that 8MG is a known DHPS binder,⁹ it was decided to measure the binding of our 8MG analogues to DHPS from *E. coli* (Figure 10). SPR data showed that compounds **3** and **4** bind DHPS with a K_D of ~4.0 μ M. Furthermore, the best HPPK binder, compound **11**, also binds DHPS with appreciable affinity (K_D ~ 8.0 μ M). These K_D values are notably better than that of 8MG (76.1 μ M). As for HPPK binding, the para OMe group in compound **3** was found to have little impact on DHPS binding affinity, and fluorine substitution at the ortho position of the benzyl group appears to be beneficial.

CONCLUSIONS

Building on our previous structure-based approaches toward inhibiting HPPK, we have used a combination of biophysical methods, including SPR, NMR spectroscopy, and X-ray crystallography, to reveal mechanistically important structural changes accompanying binding of a series of 8MG-derived substrate-site inhibitors to cofactor analogue-bound HPPK. In combination with chemical synthesis, this has resulted in the development of an advanced new lead compound (**11**) displaying an affinity for SaHPPK over 20 times greater than the previously reported parent compound, 8MG. Active site structural details for the complexes presented here will assist in the design and development of species-selective or broad-spectrum inhibitors of HPPK. In this regard, it is notable that the binding of 8MG and other analogues to EcHPPK is apparently much weaker than that to SaHPPK.⁴¹ Sequence-related structural differences, discussed here, may present avenues to increased selectivity and potency. A number of the 8MG analogues exhibit appreciable affinity for DHPS, highlighting the potential for this class of compound to be developed into dual-target inhibitors. A major focus of our future work will be on developing analogues of **11** that are able to permeate bacterial membranes, with the goal of achieving antibacterial activity.

■ EXPERIMENTAL SECTION

Samples of SaHPPK for NMR Spectroscopy. Isotopically labeled protein samples for NMR spectroscopy were prepared as described previously.²¹ *E. coli* BL21 (DE3) cells were transformed and grown overnight in 3 mL of 2× YT medium supplemented with 100 $\mu\text{g mL}^{-1}$ kanamycin for selection. The overnight culture was subcultured into 50 mL of minimal media that was grown to an OD_{600} of 0.5–0.7. This was then added to 1 L of minimal media supplemented with 1.5 g of ^{15}N ammonium chloride and/or 3 g of ^{13}C glucose and grown at 310 K until the OD_{600} was 0.5–0.8. Isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to a final optimized concentration of 0.5 mM, and expression was carried out at 293 K for 12 h. Purification was carried out as reported previously.⁴⁴

Preparation of DHPS from *E. coli*. A pET28a plasmid containing the synthesized EcDHPS sequence (Geneart) was cloned with an N-terminal hexahistidine tag and a thrombin cleavage site. *E. coli* BL21 (DE3) cells transformed with the plasmid were grown overnight in 20 mL 2× YT media supplemented with 50 $\mu\text{g mL}^{-1}$ kanamycin for selection. The overnight culture was then subcultured into fresh 2× YT (0.5L) with growth at 310 K for ~2 h until an OD_{600} of 0.5–0.7 was reached. IPTG was added to a final concentration of 0.3 mM, with expression occurring at 301 K for 22 h. The cultures were centrifuged at 5000 rpm for 10 min, and the cells were resuspended in 50 mL 50 mM Tris, pH 8.5, 5% glycerol, 5 mM MgCl_2 . An EDTA-free complete protease-inhibitor cocktail tablet (Roche) was added together with lysozyme and DNase to a final concentration of 0.4 and 0.6 mg mL^{-1} , respectively. After 10 min, the cells were sonicated, and the cell debris was removed by centrifugation at 18 000 rpm at 277 K for 30 min. The supernatant was filtered (0.45 μm filter) and loaded onto a Ni-NTA IMAC column (Qiagen). Unbound protein was washed off with 10 mM imidazole in 50 mM Tris buffer, pH 8.5, 0.1 M NaCl, 5% glycerol, 2 mM MgCl_2 , 1 mM DTT. Protein was eluted from the column with a 500 mM imidazole, DTT-free variant of the above buffer. The protein was further purified using a Superdex 75 size-exclusion 16/60 column (GE Healthcare) and eluted with 50 mM Tris, pH 8.5, 2 mM MgCl_2 , 2 mM DTT, followed by the use of a MonoQ ion-exchange 16/10 column (GE Healthcare). The column was equilibrated with 50 mM Tris, pH 8.5, 2 mM MgCl_2 , 2 mM DTT, with elution of protein using an equivalent buffer with the addition of 0.25 M NaCl. Fractions were analyzed using a 15% SDS-PAGE gel with Coomassie staining. Protein was pooled and concentrated to 2 mg mL^{-1} using a 3 kDa molecular weight cutoff ultrafiltration centrifugal device (Amicon). All samples were snap-frozen and stored at 193 K.

NMR Spectroscopy. All NMR experiments were recorded at 295 K on a Bruker Avance 600 MHz NMR spectrometer equipped with a cryoprobe and Z-axis gradient. Triple-resonance NMR experiments were performed on a sample of ~0.25 mM $^{15}\text{N}/^{13}\text{C}$ -labeled SaHPPK dissolved in a 90%/10% $\text{H}_2\text{O}/\text{D}_2\text{O}$ HEPES buffer, 1% sorbitol, and 2% $\text{DMSO}-d_6$ at pH 8.0 in the presence of 10 mM MgCl_2 and saturating amounts of AMPCPP (0.5 mM). Titrations of compound 3 or 11 to saturation was performed from a 50 mM stock dissolved in $\text{DMSO}-d_6$. Backbone assignments were obtained using the HNCA and HN(CO)CA experiments, and assignments were further confirmed using a 3D ^{15}N -edited NOESY experiment recorded with a mixing time of 120 ms. 3D experiments used a WATERGATE sequence for solvent suppression. ^{15}N heteronuclear NOE spectra were recorded on a ~0.36 mM ^{15}N -labeled sample of SaHPPK in the presence of 1 mM AMPCPP and either 600 μM 8MG or ~400 μM 3 using gradients for coherence selection and sensitivity enhancement. Three seconds of saturation was applied using a binomial train of pulses separated by a delay of 5 ms to generate the desired heteronuclear NOE and was applied off- and on-resonance in an interleaved manner, in addition to 1 s of relaxation delay. Errors were calculated from the baseline noise level. Spectra were processed using NMRPipe⁴⁵ and analyzed with XEASY⁴⁶ or SPARKY.⁴⁷ 2D ^{15}N HSQC and ^{15}N NOE experiments were typically acquired with $t_{1\text{max}}$ (^{15}N) = 51–62 ms and $t_{2\text{max}}$ (^1H) = 142 ms, whereas triple-resonance experiments were acquired with $t_{1\text{max}}$ (^{15}N) = 23.3 ms, $t_{2\text{max}}$ (^{13}C) = 10.4 ms, $t_{2\text{max}}$ (^1H) = 15.1 ms, and $t_{3\text{max}}$ (^1H) = 142 ms.

Crystallization and X-ray Structure Determination. Crystallization experiments were performed as described previously.⁴⁴ In brief, co-crystallization was set up in the C3 screens (CSIRO) at 281 K using sitting-drop vapor-diffusion method with droplets consisting of 150 nL of protein solution and 150 nL of reservoir solution and using a reservoir volume of 50 μL . Crystals of the SaHPPK in complex with AMPCPP were obtained from a solution containing 120 mM magnesium acetate, 12.6% (w/v) PEG 8000, 120 mM Tris, pH 8.5, 1 mM AMPCPP, and the protein at a concentration of 6.9 mg mL^{-1} . Crystals of the 3/AMPCPP/SaHPPK complex grew from a solution containing 275 mM ammonium nitrate and 22.1% PEG 4000 with a protein concentration of 7.5 mg mL^{-1} . Crystals of the 7/AMPCPP/SaHPPK complex were grown under similar conditions: 210 mM ammonium nitrate, 22.2% PEG 3350, and a protein concentration of 6.9 mg mL^{-1} . Data were collected at the MX-2 beamline of the Australian Synchrotron (see Table 1 for statistics) using an ADSC Quantum 315 detector, with 270 frames obtained with a one-degree oscillation angle for a complete data set. These data were indexed using XDS⁴⁸ and scaled using SCALA.⁴⁹ The SaHPPK structure (4AD6) was used to solve the initial phases of the binary and ternary complexes by molecular replacement using Phaser.⁵⁰ Refinement was performed using REFMACS,⁵¹ and the electron density maps were visualized in Coot.⁵² After several rounds of manual rebuilding, ligands and water molecules were added, and the models further refined to a resolution of 2.7 Å (R_{free} (%) = 24.2; R_{work} (%) = 19.4) for the AMPCPP complex; to a resolution of 2.0 Å (R_{free} (%) = 20.1; R_{work} (%) = 16.1) for the 3/AMPCPP/SaHPPK ternary complex, and to a resolution of 1.6 Å (R_{free} (%) = 17.8; R_{work} (%) = 16.0) for the 7/AMPCPP/SaHPPK ternary complex.

The coordinates of SaHPPK in complex with AMPCPP, in complex with 3/AMPCPP, and in complex with 7/AMPCPP have been deposited at the Protein Data Bank with accession numbers 4CYU, 4CRJ, and 4CWB, respectively.

Surface Plasmon Resonance (SPR). All SPR experiments were performed using Biacore T200 biosensor (GE Healthcare). Immobilizations were performed in HBS-EP+ running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 50 μM EDTA, 0.05% [v/v] Tween-20) at 298 K with a constant flow-rate of 10 $\mu\text{L min}^{-1}$. SaHPPK and EcDHPS proteins were covalently coupled to the NTA chip (GE Healthcare) surface using a previously described method.⁵³ Briefly, a single flow cell on the chip surface was sequentially activated by injecting (1) 40 μL of nickel sulfate and (2) 70 μL of a 1:1 mixture of NHS/EDC (*N*-hydroxysuccinimide/*N*-ethyl-*N'*-(3-diethylaminopropyl)carbodiimide). Recombinant protein was diluted in the running buffer (SaHPPK to 225 $\mu\text{g mL}^{-1}$; EcDHPS to 80 $\mu\text{g mL}^{-1}$) and injected over an activated flow cell for 20 min (200 μL). Amine-coupled surface was subsequently blocked with 70 μL of 1 M ethanolamine, pH 8.0, and then further regenerated with two 10 μL injections of 350 mM EDTA prepared in running buffer. Using this coupling approach, average immobilization levels achieved were 5400 RU for SaHPPK and 7200 RU for EcDHPS. Additionally, ubiquitin-specific-processing protease 7 (USP7) was coupled in a similar fashion to provide for an unrelated negative control surface (6600 RU). All SPR binding experiments were performed at 293 K in SPR binding buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 5 mM DTT, 10 mM MgCl_2 , 0.05% [v/v] Tween-20, 5% [v/v] DMSO). Analytes were serially diluted (3-fold) in SPR binding buffer, injected for 30 s contact time at 60 $\mu\text{L min}^{-1}$, and then allowed to dissociate for 60 s. Each analyte titration was performed in duplicate or greater. Binding sensorgrams were processed, solvent-corrected, and double-referenced using Scrubber software (BioLogic Software, Australia). SPR binding analysis of several of the compounds investigated in this study revealed dissociation rates that were not sufficiently slow to allow global fitting to a kinetic binding model, for which the k_d (dissociation rate constant) must typically be $<0.5 \text{ s}^{-1}$ for SPR instruments to be able to capture sufficient data points during the dissociation phase. Therefore, to determine binding affinities (K_D values), responses at equilibrium for each analyte were fitted to a 1:1 steady-state affinity model available within Scrubber using 8MG as a reference, as previously described.²⁰

KinaseGlo Biochemical Assay. HPPK activity was quantified using a KinaseGlo assay kit (Promega) as previously reported.²¹ In this, firefly luciferase utilizes the remaining ATP after HPPK catalysis, producing a

luminescence signal that is directly proportional to ATP concentration. The enzyme activity and optimum concentration to define kinetic parameters were optimized as described previously.²¹ For kinetic measurements, an optimized HPPK concentration of 0.4 ng μL^{-1} assay volume was used, which allowed for monitoring of the first 10% of reaction turnover within a reasonable time period (20 min). Measurements were performed in 96-well plates using assay buffer (100 mM Tris-HCl, pH 8.5, 10 mM MgCl_2 , 0.01% (w/v) BSA, 0.01% (v/v) Tween 20 and 10 mM β -mercaptoethanol). Typically, 5 μL of test compound (dissolved in 50% DMSO) and 20 μL of enzyme were added to each well followed by 25 μL of assay buffer, giving 0.3 μM pterin and 0.2 μM ATP in a total reaction volume of 50 μL . After 20 min incubation at room temperature, the enzymatic reaction was stopped with 50 μL of KinaseGlo reagent. Luminescence was recorded after a further 10 min using a FLUOstar Optima plate reader (BMG, Labtech Ltd.). Reactions were performed in triplicate. Kinetic data and inhibition data were fit to Michaelis–Menten and sigmoidal dose–response equations, respectively, using GraphPad Prism.

Molecular Modeling. Molecular modeling was performed using the Schrödinger Suite 2014 (www.schrodinger.com) through the Maestro interface (Maestro, version 9.7).⁵⁴ Protein preparation of 4CWB was performed with the Protein Preparation Wizard workflow implemented by Schrödinger (Epik, version 2.7),⁵⁵ with deletion of all waters. In order to eliminate any bond length or bond angle biases in the structures, compound 11 was subjected to a full minimization prior to docking using LigPrep (LigPrep, version 2.9).⁵⁶ Docking was carried out with Glide,⁵⁷ version 6.2, using Extra Precision (XP) mode.

Compound Procurement and Analysis. 8MG and compound 3 were purchased from Sigma-Aldrich and TimTec, respectively. All other compounds were synthesized as described below. In all cases, ^1H NMR spectra were recorded on a 400 MHz Bruker NMR spectrometer, and chemical shifts were referenced to the solvent peak. Analytical reversed-phase high-performance liquid chromatography (RP-HPLC) was performed on an Agilent 1260 Infinity HPLC system using an Agilent Eclipse Plus C18 column (100 \times 4.6 mm i.d., 3.5 μm) with a flow rate of 1 mL min^{-1} and UV detection at 214 and 254 nm. Elution was achieved with standard HPLC buffers (buffer A: 99.9% H_2O /0.1% TFA; buffer B: 99.9% CH_3CN /0.1% TFA) using a gradient from 5% B/95% A to 100% B over 10 min. All compounds were determined to be >95% purity by this method.

General Procedure A for the Synthesis of Compounds 4–7 and 9. 8-Mercaptoguanine (0.200 g, 1.09 mmol) was dissolved in 0.4 M NaOH (5.5 mL), and to the solution was added phenacyl bromide (or analogue) (0.24 g, 1.2 mmol) in ethanol (0.9 mL). The reaction was allowed to stir for 2 h, following which a white precipitate formed in solution, which was collected by vacuum filtration to give the title compound as a white amorphous solid.

2-Amino-8-((2-oxo-2-phenylethyl)thio)-1,9-dihydro-6H-purin-6-one (4). Compound 4 was synthesized using general procedure A. Yield 50%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 12.51 (bs, 1H), 10.90 (bs, 1H), 8.03–8.01 (m, 2H), 7.67 (t, J = 7.4 Hz, 1H), 7.55 (t, J = 7.7 Hz, 2H), 6.45 (s, 2H), 4.86 (s, 2H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{13}\text{H}_{11}\text{N}_5\text{O}_2\text{S}$, 302.0706; found, 302.0710.

2-Amino-8-((1-oxo-1-phenylpropan-2-yl)thio)-1,9-dihydro-6H-purin-6-one (5). Compound 5 was synthesized using general procedure A including the following work up. The reaction mixture was diluted with 0.5 M NaOH (7 mL). The aqueous layer was extracted with EtOAc (3 \times 3 mL) and acidified with acetic acid (2 mL). The precipitate solid was collected by filtration and dried to give solid product. Yield 28%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 10.91 (bs, 1H), 8.01 (d, J = 7.4 Hz, 2H), 7.65 (t, J = 7.3 Hz, 1H), 7.52 (t, J = 7.7 Hz, 2H), 6.46 (bs, 2H), 5.43 (q, J = 6.7 Hz, 1H), 1.53 (d, J = 6.9 Hz, 3H). HRMS: m/z calcd for $[\text{M} + \text{Na}]^+$ $\text{C}_{14}\text{H}_{13}\text{N}_5\text{O}_2\text{S}$, 338.0682; found, 338.0688.

2-Amino-8-((2-(2-methoxyphenyl)-2-oxoethyl)thio)-1,9-dihydro-6H-purin-6-one (6). Compound 6 was synthesized using general procedure A including the following work up. The reaction mixture was diluted with 0.5 M NaOH (7 mL). The aqueous layer was extracted with EtOAc (3 \times 3 mL) and acidified with acetic acid (2 mL). The precipitate solid was collected by filtration and dried to give solid product. Yield 63%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 10.58 (bs, 1H),

7.64 (dd, J = 7.7, 1.8 Hz, 1H), 7.59 (m, 1H), 7.21 (d, J = 8.3 Hz, 1H), 7.08–7.01 (m, 1H), 6.28 (bs, 2H), 4.67 (s, 2H), 3.92 (s, 3H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{14}\text{H}_{13}\text{N}_5\text{O}_3\text{S}$, 332.0812; found, 332.0821.

8-((2-([1,1'-Biphenyl]-4-yl)-2-oxoethyl)thio)-2-amino-1,9-dihydro-6H-purin-6-one (7). Compound 7 was synthesized using general procedure A. Yield 9%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 10.97 (s, 1H), 8.11 (d, J = 8.5 Hz, 2H), 7.82 (d, J = 8.5 Hz, 2H), 7.77–7.74 (m, 2H), 7.50 (m, 2H), 7.45–7.43 (m, 1H), 6.39 (s, 2H), 4.80 (s, 2H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{19}\text{H}_{15}\text{N}_5\text{O}_2\text{S}$, 378.1019; found, 378.1028.

2-Amino-8-((2-oxopropyl)thio)-1,9-dihydro-6H-purin-6-one (9). Compound 9 was synthesized using general procedure A using 4 equiv of chloroacetone and 8 mL of 0.4 M NaOH. Yield 55%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 12.60 (bs, 1H), 10.65 (bs, 1H), 6.33 (s, 2H), 4.16 (s, 2H), 2.23 (s, 3H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_8\text{H}_9\text{N}_5\text{O}_2\text{S}$, 240.0550; found, 240.0553.

General Procedure B for the Synthesis of Compounds 8 and 10–13. 8-Mercaptoguanine (0.10 g, 0.55 mmol) was dissolved in 0.5 M NaOH (3 mL), and to the solution was added substituted benzyl or phenethyl bromide (0.61 mmol). The mixture was stirred at room temperature for 5–24 h and then diluted with 0.5 M NaOH (8 mL). The aqueous layer was extracted with EtOAc (3 \times 20 mL), acidified with acetic acid (pH 5.0), and stirred for 15–20 min. The precipitated solid was collected by filtration, washed thoroughly with water and ethanol, and dried to give the solid product.

2-Amino-8-((4-methoxyphenethyl)thio)-1,9-dihydro-6H-purin-6-one (8). Compound 8 was synthesized using general procedure B. Yield 73%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 12.5 (bs, 1H), 10.5 (bs, 1H), 7.18–6.84 (m, 4H), 6.55 (s, 2H), 3.72 (s, 3H), 3.33 (t, J = 7.7 Hz, 2H), 2.88 (t, J = 7.1 Hz). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{14}\text{H}_{15}\text{F}_3\text{N}_5\text{O}_2\text{S}$, 318.1019; found, 318.1021.

2-Amino-8-(benzylthio)-1,9-dihydro-6H-purin-6-one (10). Compound 10 was synthesized using general procedure B. Yield 36%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 12.5 (bs, 1H, NH), 10.5 (bs, 1H), 7.29–7.21 (m, 5H), 6.57 (s, 2H), 4.34 (s, 2H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{12}\text{H}_{11}\text{N}_5\text{OS}$, 274.0757; found, 274.0760.

2-Amino-8-((2-fluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (11). Compound 11 was synthesized using general procedure B. Yield 66%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 12.5 (bs, 1H), 10.5 (bs, 1H), 7.37–7.03 (m, 4H), 6.30 (s, 2H), 4.30 (s, 2H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{12}\text{H}_{10}\text{FN}_5\text{OS}$, 292.0663; found, 292.0666.

2-Amino-8-((2-(trifluoromethyl)benzyl)thio)-1,9-dihydro-6H-purin-6-one (12). Compound 12 was synthesized using general procedure B. Yield 52%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 12.6 (bs, 1H), 10.6 (bs, 1H), 7.75–7.48 (m, 4H), 6.36 (s, 2H), 4.53 (s, 2H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{13}\text{H}_{10}\text{F}_3\text{N}_5\text{OS}$, 342.0631; found, 342.0636.

4-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methylbenzonitrile (13). Compound 13 was synthesized using general procedure B. Yield 73%. ^1H NMR (DMSO- d_6 , 400 MHz): δ 10.89 (bs, 1H), 7.77 (d, J = 8.3 Hz, 2H), 7.55 (d, J = 8.3 Hz, 2H), 6.43 (bs, 2H), 4.45 (s, 2H). HRMS: m/z calcd for $[\text{M} + \text{H}]^+$ $\text{C}_{13}\text{H}_{10}\text{N}_6\text{OS}$, 299.0710; found, 299.0699.

■ ASSOCIATED CONTENT

Supporting Information

CSP data for the binding of 11 to AMPCPP saturated SaHPPK. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Accession Codes

The coordinates of SaHPPK in complex with AMPCPP, in complex with 3/AMPCPP, and in complex with 7/AMPCPP have been deposited at the Protein Data Bank with accession numbers 4CYU, 4CRJ, and 4CWB, respectively.

■ AUTHOR INFORMATION

Corresponding Author

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

All SaHPPK crystals were grown at the C3 Crystallization Centre at CSIRO, Parkville, Australia, and X-ray data were obtained at the MX2 beamline at the Australian Synchrotron, Victoria, Australia. We thank OpenEye for a license to use their software. J.B.B. was supported by an NHMRC Research Fellowship (1020411).

■ ABBREVIATIONS USED

HMDP, 6-hydroxymethyl-7,8-dihydro-pterin; SMX, sulfamethoxazole; TMP, trimethoprim; SaHPPK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *S. aureus*; EcHPPK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *E. coli*; EcDHPS, dihydropteroate synthase from *E. coli*; MRSA, methicillin-resistant *S. aureus*; 8MG, 8-mercaptopuanine; SPR, surface plasmon resonance; HSQC, heteronuclear single quantum coherence; SAR, structure–activity relationships

■ REFERENCES

- (1) Bermingham, A.; Derrick, J. P. The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. *BioEssays* **2002**, *24*, 637–648.
- (2) Swarbrick, J. D.; Iliades, P.; Simpson, J. S.; Macreadie, I. Folate biosynthesis - reappraisal of old and novel targets in the search for new antimicrobials. *Open Enzyme Inhib. J.* **2008**, *1*, 12–33.
- (3) Masters, P. A.; O'Bryan, T. A.; Zurlo, J.; Miller, D. Q.; Joshi, N. Trimethoprim-sulfamethoxazole revisited. *Arch. Int. Med.* **2003**, *163*, 402–410.
- (4) Adra, M.; Lawrence, K. R. Trimethoprim/sulfamethoxazole for treatment of severe *Staphylococcus aureus* infections. *Ann. Pharmacother.* **2004**, *38*, 338–341.
- (5) Forgacs, P.; Wengenack, N. L.; Hall, L.; Zimmerman, S. K.; Silverman, M. L.; Roberts, G. D. Tuberculosis and trimethoprim-sulfamethoxazole. *Antimicrob. Agents Chemother.* **2009**, *53*, 4789–4793.
- (6) Martin, S. I.; Fishman, J. A. *Pneumocystis pneumonia* in solid organ transplant recipients. *Am. J. Transplant.* **2009**, *9*, S227–S233.
- (7) Pashley, T. V.; Volpe, F.; Pudney, M.; Hyde, J. E.; Sims, P. F.; Delves, C. J. Isolation and molecular characterization of the bifunctional hydroxymethyl-dihydropterin pyrophosphokinase-dihydropteroate synthase gene from *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* **1997**, *86*, 37–47.
- (8) Thera, M. A.; Sehdev, P. S.; Coulibaly, D.; Traore, K.; Garba, M. N.; Cissoko, Y.; Kone, A.; Guindo, A.; Dicko, A.; Beavogui, A. H.; Djimde, A. A.; Lyke, K. E.; Diallo, D. A.; Doumbo, O. K.; Plowe, C. V. Impact of trimethoprim-sulfamethoxazole prophylaxis on falciparum malaria infection and disease. *J. Infect. Dis.* **2005**, *192*, 1823–1829.
- (9) Hevener, K. E.; Yun, M. K.; Qi, J.; Kerr, I. D.; Babaoglu, K.; Hurdle, J. G.; Balakrishna, K.; White, S. W.; Lee, R. E. Structural studies of pterin-based inhibitors of dihydropteroate synthase. *J. Med. Chem.* **2010**, *53*, 166–177.
- (10) Frey, K. M.; Liu, J.; Lombardo, M. N.; Bolstad, D. B.; Wright, D. L.; Anderson, A. C. Crystal structures of wild-type and mutant methicillin-resistant *Staphylococcus aureus* dihydrofolate reductase reveal an alternate conformation of NADPH that may be linked to trimethoprim resistance. *J. Mol. Biol.* **2009**, *387*, 1298–1308.
- (11) Bourne, C. R.; Bunce, R. A.; Bourne, P. C.; Berlin, K. D.; Barrow, E. W.; Barrow, W. W. Crystal structure of *Bacillus anthracis* dihydrofolate reductase with the dihydrophthalazine-based trimethoprim derivative RAB1 provides a structural explanation of potency and selectivity. *Antimicrob. Agents Chemother.* **2009**, *53*, 3065–3073.
- (12) Yun, M. K.; Wu, Y.; Li, Z.; Zhao, Y.; Waddell, M. B.; Ferreira, A. M.; Lee, R. E.; Bashford, D.; White, S. W. Catalysis and sulfa drug resistance in dihydropteroate synthase. *Science* **2012**, *335*, 1110–1114.
- (13) Zhao, Y.; Hammoudeh, D.; Yun, M. K.; Qi, J.; White, S. W.; Lee, R. E. Structure-based design of novel pyrimido[4,5-c]pyridazine derivatives as dihydropteroate synthase inhibitors with increased affinity. *ChemMedChem.* **2012**, *7*, 861–870.
- (14) Hammoudeh, D. I.; Date, M.; Yun, M. K.; Zhang, W.; Boyd, V. A.; Viacava Follis, A.; Griffith, E.; Lee, R. E.; Bashford, D.; White, S. W. Identification and characterization of an allosteric inhibitory site on dihydropteroate synthase. *ACS Chem. Biol.* **2014**, 1294–1302.
- (15) Haruki, H.; Pedersen, M. G.; Gorska, K. I.; Pojer, F.; Johnsson, K. Tetrahydrobiopterin biosynthesis as an off-target of sulfa drugs. *Science* **2013**, *340*, 987–91.
- (16) Kluytmans-Vandenbergh, M. F.; Kluytmans, J. A. Community-acquired methicillin-resistant *Staphylococcus aureus*: current perspectives. *Clin. Microbiol. Infect.* **2006**, *12*, 9–15.
- (17) Kollef, M. H.; Micek, S. T. Methicillin-resistant *Staphylococcus aureus*: a new community-acquired pathogen? *Curr. Opin. Infect. Dis.* **2006**, *19*, 161–168.
- (18) Ohlsen, K.; Lorenz, U. Novel targets for antibiotics in *Staphylococcus aureus*. *Future Microbiol.* **2007**, *2*, 655–666.
- (19) Spaulding, A. R.; Salgado-Pabon, W.; Merriman, J. A.; Stach, C. S.; Ji, Y.; Gillman, A. N.; Peterson, M. L.; Schlievert, P. M. Vaccination against *Staphylococcus aureus* pneumonia. *J. Infect. Dis.* **2014**, *209*, 1955–1962.
- (20) Chhabra, S.; Barlow, N.; Dolezal, O.; Hattarki, M. K.; Newman, J.; Peat, T. S.; Graham, B.; Swarbrick, J. D. Exploring the chemical space around 8-mercaptopuanine as a route to new inhibitors of the folate biosynthesis enzyme HPPK. *PLoS One* **2013**, *8*, e59535.
- (21) Chhabra, S.; Dolezal, O.; Collins, B. M.; Newman, J.; Simpson, J. S.; Macreadie, I. G.; Fernley, R.; Peat, T. S.; Swarbrick, J. D. Structure of *S. aureus* HPPK and the discovery of a new substrate site inhibitor. *PLoS One* **2012**, *7*, e29444.
- (22) Derrick, J. P. The structure and mechanism of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Vitam. Horm.* **2008**, *79*, 411–433.
- (23) Blaszczyk, J.; Shi, G.; Li, Y.; Yan, H.; Ji, X. Reaction trajectory of pyrophosphoryl transfer catalyzed by 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Structure* **2004**, *12*, 467–475.
- (24) Xiao, B.; Shi, G.; Gao, J.; Blaszczyk, J.; Liu, Q.; Ji, X.; Yan, H. Unusual conformational changes in 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase as revealed by X-ray crystallography and NMR. *J. Biol. Chem.* **2001**, *276*, 40274–40281.
- (25) Garcon, A.; Levy, C.; Derrick, J. P. Crystal structure of the bifunctional dihydroneopterin aldolase/6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Streptococcus pneumoniae*. *J. Mol. Biol.* **2006**, *360*, 644–653.
- (26) Hennig, M.; Dale, G. E.; D'Arcy, A.; Danel, F.; Fischer, S.; Gray, C. P.; Jolidon, S.; Muller, F.; Page, M. G.; Pattison, P.; Oefner, C. The structure and function of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Haemophilus influenzae*. *J. Mol. Biol.* **1999**, *287*, 211–219.
- (27) Lawrence, M. C.; Iliades, P.; Fernley, R. T.; Berglez, J.; Pilling, P. A.; Macreadie, I. G. The three-dimensional structure of the bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/dihydropteroate synthase of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **2005**, *348*, 655–670.
- (28) Xiao, B.; Shi, G.; Chen, X.; Yan, H.; Ji, X. Crystal structure of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, a potential target for the development of novel antimicrobial agents. *Structure* **1999**, *7*, 489–496.
- (29) Wood, H. C. S. Specific inhibition of dihydrofolate biosynthesis—a new approach to chemotherapy. In *Chemistry and Biology of Pteridines*; Pfeleiderer, W., Ed.; Walter de Gruyter: New York, 1975.
- (30) Shi, G.; Blaszczyk, J.; Ji, X.; Yan, H. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: synthesis and biochemical and crystallographic studies. *J. Med. Chem.* **2001**, *44*, 1364–1371.
- (31) Shi, G.; Shaw, G.; Li, Y.; Wu, Y.; Yan, H.; Ji, X. Bisubstrate analog inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase:

new lead exhibits a distinct binding mode. *Bioorg. Med. Chem.* **2012**, *20*, 4303–4309.

(32) Shi, G.; Shaw, G.; Liang, Y. H.; Subburaman, P.; Li, Y.; Wu, Y.; Yan, H.; Ji, X. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: new design with improved properties. *Bioorg. Med. Chem.* **2012**, *20*, 47–57.

(33) Bourne, C. R. Utility of the biosynthetic folate pathway for targets in antimicrobial discovery. *Antibiotics* **2014**, *3*, 28.

(34) Stammers, D. K.; Achari, A.; Somers, D. O.; Bryant, P. K.; Rosemond, J.; Scott, D. L.; Champness, J. N. 2.0 Å X-ray structure of the ternary complex of 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase from *Escherichia coli* with ATP and a substrate analogue. *FEBS Lett.* **1999**, *456*, 49–53.

(35) Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(36) Medek, A.; Hajduk, P. J.; Mack, J.; Fesik, S. W. The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands. *J. Am. Chem. Soc.* **2000**, *122*, 1241–1242.

(37) Lescop, E.; Lu, Z.; Liu, Q.; Xu, H.; Li, G.; Xia, B.; Yan, H.; Jin, C. Dynamics of the conformational transitions in the assembling of the Michaelis complex of a bisubstrate enzyme: a ¹⁵N relaxation study of *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Biochemistry* **2009**, *48*, 302–312.

(38) Yang, R.; Lee, M. C.; Yan, H.; Duan, Y. Loop conformation and dynamics of the *Escherichia coli* HPPK apo-enzyme and its binary complex with MgATP. *Biophys. J.* **2005**, *89*, 95–106.

(39) Iwaoka, M.; Isozumi, N. Hypervalent nonbonded interactions of a divalent sulfur atom. Implications in protein architecture and the functions. *Molecules* **2012**, *17*, 7266–7283.

(40) Blaszczyk, J.; Shi, G.; Yan, H.; Ji, X. Catalytic center assembly of HPPK as revealed by the crystal structure of a ternary complex at 1.25 Å resolution. *Structure* **2000**, *8*, 1049–1058.

(41) Yun, M. K.; Hoagland, D.; Kumar, G.; Waddell, M. B.; Rock, C. O.; Lee, R. E.; White, S. W. The identification, analysis and structure-based development of novel inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Bioorg. Med. Chem.* **2014**, *22*, 2157–2165.

(42) Zhou, P.; Zou, J. W.; Tian, F. F.; Shang, Z. C. Fluorine bonding - how does it work in protein-ligand interactions? *J. Chem. Inf. Model.* **2009**, *49*, 2344–2355.

(43) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.

(44) Chhabra, S.; Newman, J.; Peat, T. S.; Fernley, R. T.; Caine, J.; Simpson, J. S.; Swarbrick, J. D. Crystallization and preliminary X-ray analysis of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2010**, *66*, 575–578.

(45) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277–293.

(46) Bartels, C.; Xia, T. H.; Billeter, M.; Guntert, P.; Wuthrich, K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **1995**, *6*, 1–10.

(47) Goddard, T. D.; Kneller, D. G. SPARKY 3; University of California: San Francisco, CA.

(48) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–32.

(49) Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760–763.

(50) Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 432–438.

(51) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240–255.

(52) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(53) Kimple, A. J.; Muller, R. E.; Siderovski, D. P.; Willard, F. S. A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance. *Methods Mol. Biol.* **2010**, *627*, 91–100.

(54) *Maestro*, version 9.7; Schrödinger, LLC: New York, 2014.

(55) *Schrödinger Suite 2014-1: Protein Preparation Wizard*, *Epik* version 2.7, *Impact* version 6.2, *Prime* version 3.4; Schrödinger, LLC: New York, 2014.

(56) *LigPrep*; Schrödinger, LLC: New York, 2014.

(57) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196.

Supporting Information

Structure-based Design and Development of Functionalized Mercaptoguanine Derivatives as Inhibitors of the Folate Biosynthesis Pathway Enzyme 6-Hydroxymethyl-7,8- dihydropterin Pyrophosphokinase from *Staphylococcus aureus*.

Matthew L. Dennis, Sandeep Chhabra, Zhong-Chang Wang, Aaron Debono, Olan Dolezal, Janet Newman, Noel P. Pitcher, Raphael Rahmani, Ben Cleary, Nicholas Barlow, Meghan Hattarki, Bim Graham, Thomas S. Peat, Jonathan B. Baell, and James D. Swarbrick.*

Contents:

Figure S1.

Figure S1: Raw CSPs corresponding to the change in weighted average chemical shifts observed for **11**/AMPCPP/SaHPPK relative to those for 8MG/AMPCPP/SaHPPK in the 2D ^{15}N HSQC spectra. The pink, blue and red horizontal lines signify 1, 2 and 3 standard deviations of the CSPs, as shown in Figure 3C. The broadened Gly47 is represented by a yellow bar, and the side-chain of Trp89 by a grey bar.

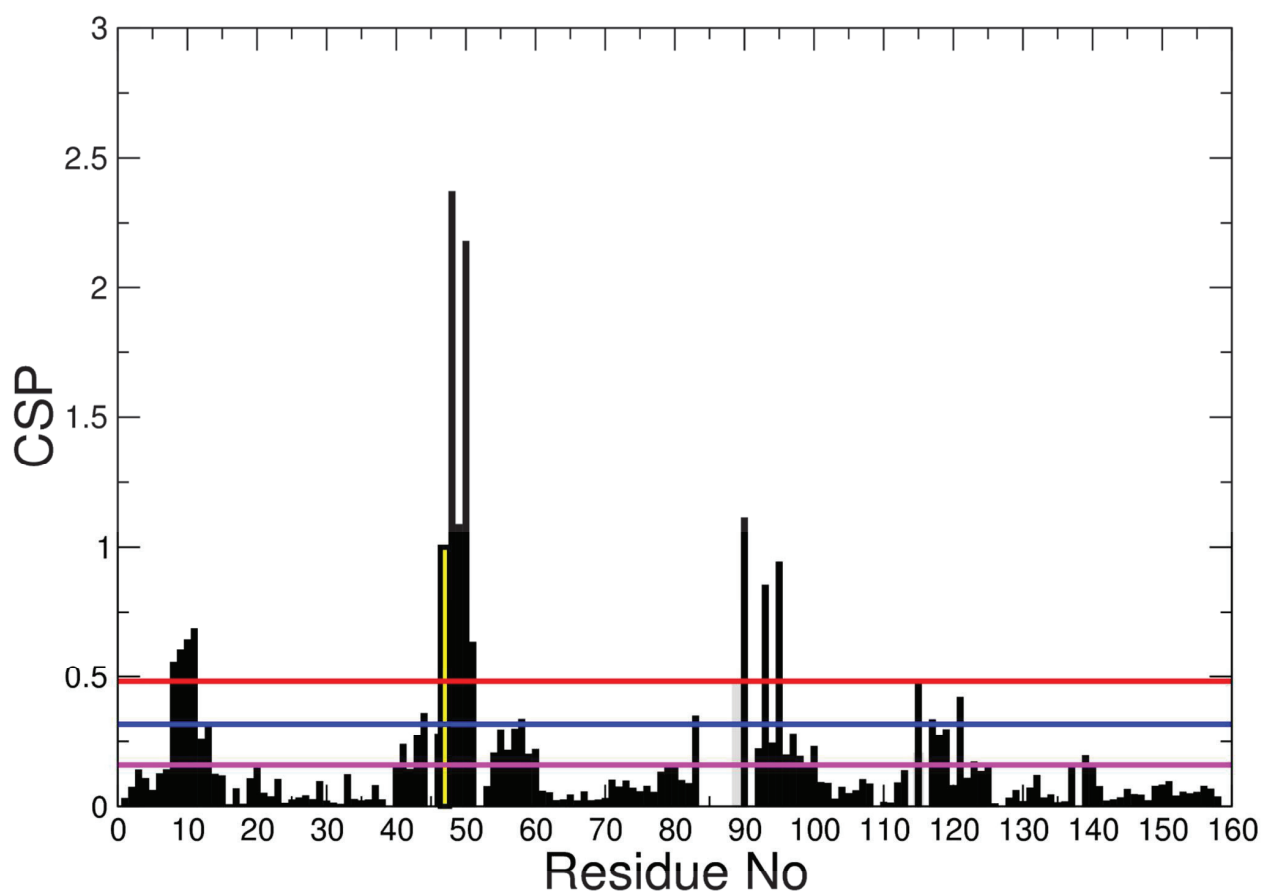
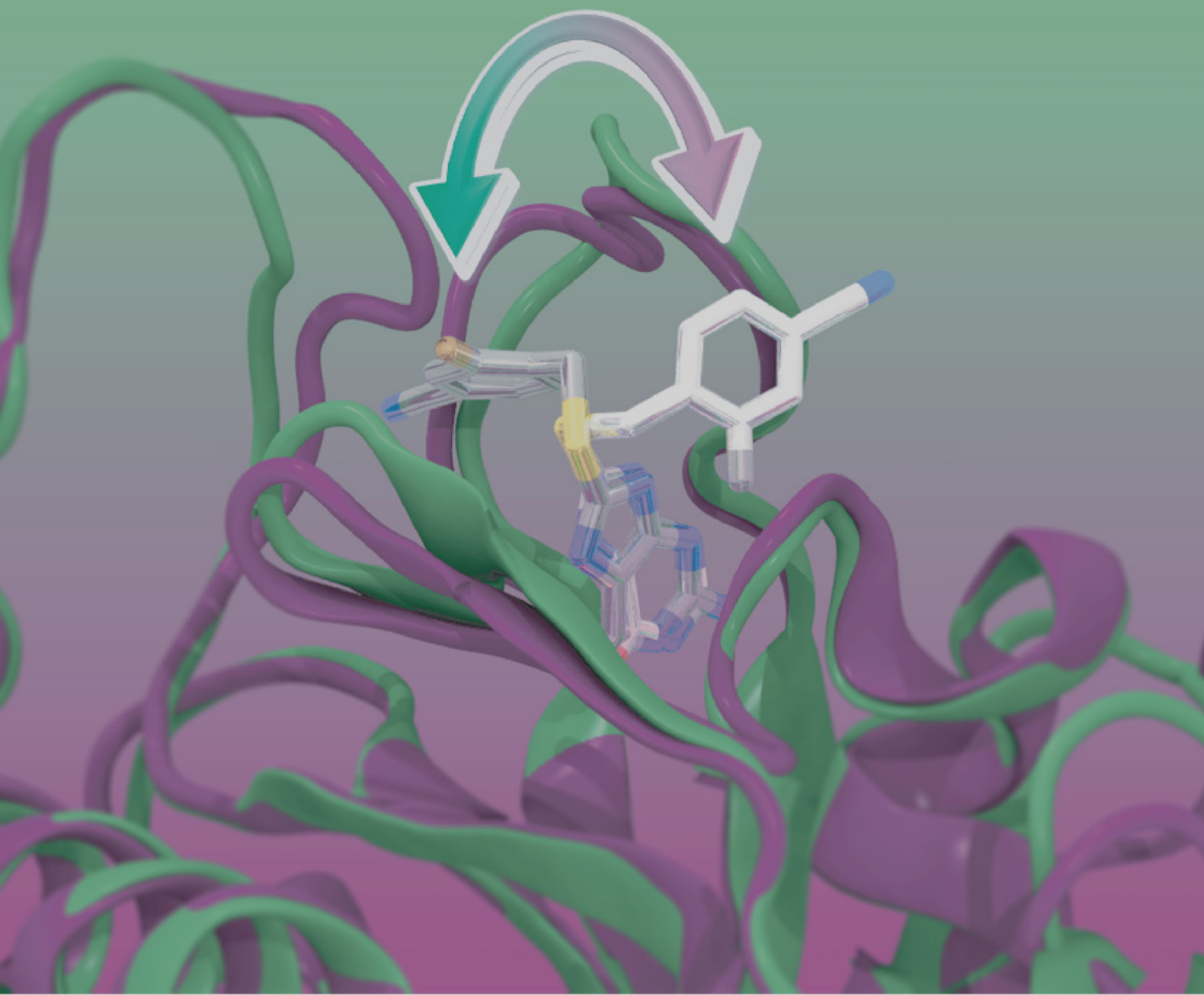


Figure S1

"Structural basis for the selective binding of inhibitors to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus* and *Escherichia coli*"

3



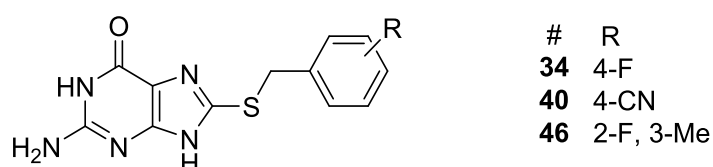
A significant (~20-fold) improvement in affinity over the parent compound, 8MG, had been achieved in SaHPPK with a novel compound. A paper exploring similar compounds, but against the *E. coli* form of the HPPK enzyme, was published while the previous manuscript was being prepared.¹ Four compounds from these papers were the same, however the affinities determined were markedly different (up to 44-fold), with an altered rank ordering. This disparity was of interest to our group and the gene for EcHPPK was purchased. Transformation, expression, and purification of EcHPPK were successful, yielding protein suitable for study via the same biophysical techniques applied to SaHPPK in the previous paper. Crystallisation trials for EcHPPK were performed with various ligands, good crystal growth being generally observed. Crystals of the EcHPPK/AMPCPP binary complex can, however, grow under the same conditions as the ligand-bound ternary complex — a fact realised only upon solving one such structure to reveal an empty pterin-binding site.

The use of UV imaging in crystallisation derives from the intrinsic fluorescence of the tryptophan residue, and is an established technique,² wherein protein crystals will glow brightly under UV light, while salt or small molecule crystals will generally appear dull. Fortuitously, the guanine moiety present in the compound series absorbs UV light. UV imaging of crystallisation drops of EcHPPK can therefore be used to reveal ligand binding. EcHPPK crystals with or without bound ligand will appear darker or brighter, respectively, than the surrounding drop. Interestingly, both types of crystal were observed to occur not only in the same drop, but in the same crystal cluster (Figure 3.1) — separating and mounting only the desired ligand-bound complex thus occasionally required appropriate “crystal surgery”.

Crystallisation of SaHPPK complexes with our lead compounds proved more difficult. A lack of crystals meant a lack of crystal conditions around which to optimise.

Crystal lattices are mediated by intermolecular interactions specific to each protein or protein-ligand complex.³ Therefore, while there are chemicals (e.g. salts, PEGs) that aid in the crystallisation of proteins in general, certain compounds will act unpredictably to create or stabilise specific intermolecular interactions and aid crystallisation. Screening a variety of compounds in the hope that one will act in this manner is thus dubbed ‘silver bullet screening’.⁴ This method was employed in an attempt to generate a new crystal form for ternary complexed SaHPPK.

Crystallisation conditions that gave a degree of precipitation were chosen as a starting point to screen for a ‘silver bullet’. 96 different compounds were tested, crystal growth was observed for the ternary complex **46**/AMPCPP/SaHPPK only in the presence of thiocyanate. A screen to optimise the concentration of thiocyanate, and a further screen to optimise the base components, was undertaken. While diffraction of **46**/AMPCPP/SaHPPK crystals remained weak, despite consistent growth, a crystal complex of **40**/AMPCPP/SaHPPK was solved to 1.96 Å in a new crystal form under the same conditions. Additional SaHPPK complexes were solved using the same conditions, however, as noted for the previous crystal form, both crystal growth and diffraction was highly ligand-dependent.



Scheme 3.1. Chemical structures of selected compounds reported in the following paper.

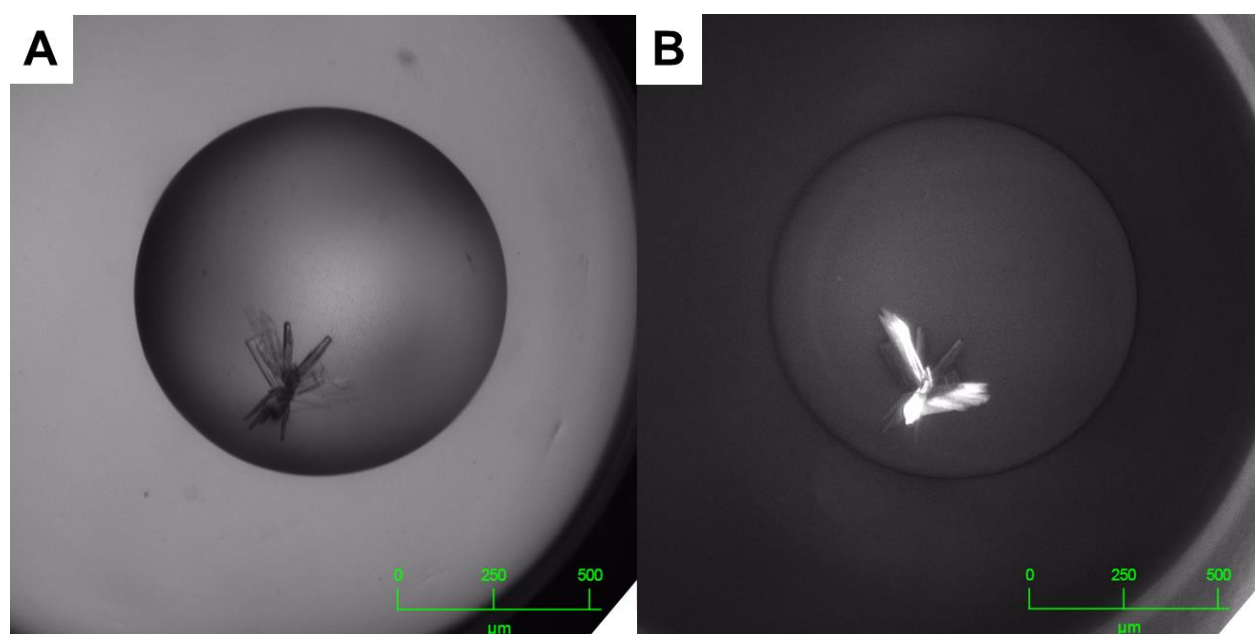


Figure 3.1. UV imaging distinguishes ligand-bound AMPCPP/EcHPPK crystals. Crystal well from which the structure of the **34**/AMPCPP/EcHPPK complex was obtained, viewed using (A) visible and (B) UV imaging.

The most significant finding from this structural data is that of a novel binding pocket, into which the benzyl pendant group projects. This pocket is both species- and ligand-specific, forming only in the SaHPPK enzyme and only with benzyl-substituted 8MG-derivatives. NMR chemical shift mapping experiments revealed two corroborating sets of CSP signatures; one for an acetophenone ligand and another for the benzyl series, which proved useful to characterise the binding site for the benzyl analogues, especially those that could not be crystallized.

The EcHPPK and SaHPPK crystal structures obtained in this publication elegantly explain the observed SAR, and provide a rational path forward to more potent and selective inhibitors of SaHPPK. If the infecting organism is known, administration of a selective antibiotic can be beneficial in avoiding the disturbance of the intestinal flora, as demonstrated recently,⁵ thereby reducing both the risk of secondary infections and antibiotic resistance.⁶⁻⁸

Readers should be aware when considering the following paper: although many compounds (particularly the lead compounds 8MG, **14** and **18**) were re-tested via SPR throughout this thesis, the uncertainties presented for K_D values indicate the standard deviation of technical replicates (i.e. discrete compound sample preparation, but performed on the same chip and therefore not independent experiments), and should be interpreted appropriately.

1. Yun, M. K.; Hoagland, D.; Kumar, G.; Waddell, M. B.; Rock, C. O.; Lee, R. E.; White, S. W. The identification, analysis and structure-based development of novel inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Bioorg. Med. Chem.* **2014**, 22, 2157–2165.
2. Desbois, S.; Seabrook, S. A.; Newman, J. Some practical guidelines for UV imaging in the protein crystallization laboratory. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2013**, 69, 201–208.
3. Derewenda, Z. Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr D Biol Crystallogr* **2010**, 66, 604–15.
4. McPherson, A.; Cudney, B. Searching for silver bullets: an alternative strategy for crystallizing macromolecules. *J Struct Biol* **2006**, 156, 387–406.
5. Yao, J.; Carter, R. A.; Vuagniaux, G.; Barbier, M.; Rosch, J. W.; Rock, C. O. A pathogen-selective antibiotic minimizes disturbance to the microbiome. *Antimicrob. Agents Chemother.* **2016**, 60, 4264–4273.
6. Fowler, S.; Webber, A.; Cooper, B. S.; Phimister, A.; Price, K.; Carter, Y.; Kibbler, C. C.; Simpson, A. J.; Stone, S. P. Successful use of feedback to improve antibiotic prescribing and reduce *Clostridium difficile* infection: a controlled interrupted time series. *J. Antimicrob. Chemother.* **2007**, 59, 990–995.
7. Yates, C. M.; Shaw, D. J.; Roe, A. J.; Woolhouse, M. E.; Amyes, S. G. Enhancement of bacterial competitive fitness by apramycin resistance plasmids from non-pathogenic *Escherichia coli*. *Biol. Lett.* **2006**, 2, 463–465.
8. Lemon, K. P.; Armitage, G. C.; Relman, D. A.; Fischbach, M. A. Microbiota-targeted therapies: an ecological perspective. *Sci. Transl. Med.* **2012**, 4, 137rv5.

Declaration for Thesis Chapter 3

Chapter 3 presents the published article co-authored with Noel P. Pitcher, Michael D. Lee, Aaron DeBono, Zhong-Chang Wang, Jitendra Harjani, Raphaël Rahmani, Ben Cleary, Meghan Hattarki, Thomas S. Peat, Jonathan B. Baell, and James D. Swarbrick.

Declaration by candidate

In the case of Chapter 3, the nature and extent of my contribution to the work was as the following:

Nature of Contribution	Extent of Contribution (%)
Protein expression and purification, NMR and X-ray crystallography experiments, SPR, biochemical assays, molecular modelling	70

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Noel P. Pitcher ¹	Synthetic chemistry, intellectual input	5
Michael D. Lee ¹	Synthetic chemistry, intellectual input	5
Aaron DeBono ¹	Synthetic chemistry, intellectual input	
Zhong-Chang Wang ³	Synthetic chemistry	
Jitendra Harjani ¹	Synthetic chemistry	
Raphaël Rahmani ¹	Supervision of chemists	
Ben Cleary ¹	Supervision of chemists	5

Thomas S. Peat ²	X-ray crystallography, intellectual input	
Jonathan B. Baell ¹	Synthetic chemistry, intellectual input	
James D. Swarbrick ¹	NMR experiments, intellectual input	

Declaration by co-authors

The undersigned hereby certify that:


- (1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors;
- (2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- (3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- (4) there are no other authors of the publication according to these criteria;
- (5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- (6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:



Location(s)

¹Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

²CSIRO Biosciences Program, Parkville, Victoria 3052, Australia

³State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, People's Republic of China

		Chapter 3 - Declaration
Signature 1		19/12/2016
Signature 2		
Signature 3		

	Date	
Signature 2		19/12/16
Signature 3		20/12/16

Structural Basis for the Selective Binding of Inhibitors to 6-Hydroxymethyl-7,8-dihydropterin Pyrophosphokinase from *Staphylococcus aureus* and *Escherichia coli*

Matthew L. Dennis,^{†,‡} Noel P. Pitcher,[†] Michael D. Lee,[†] Aaron J. DeBono,[†] Zhong-Chang Wang,^{†,§} Jitendra R. Harjani,[†] Raphaël Rahmani,[†] Ben Cleary,[†] Thomas S. Peat,[‡] Jonathan B. Baell,[†] and James D. Swarbrick^{*,†}

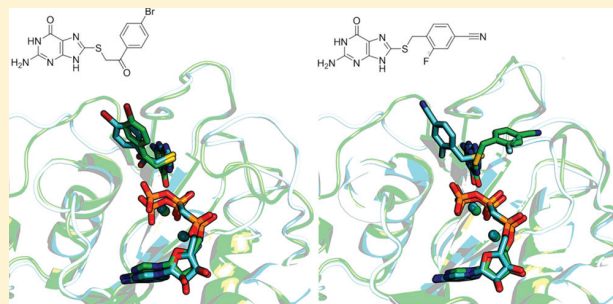
[†]Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

[‡]CSIRO Biosciences Program, Parkville, Victoria 3052, Australia

[§]State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, People's Republic of China

S Supporting Information

ABSTRACT: 6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) is a member of the folate biosynthesis pathway found in prokaryotes and lower eukaryotes that catalyzes the pyrophosphoryl transfer from the ATP cofactor to a 6-hydroxymethyl-7,8-dihydropterin substrate. We report the chemical synthesis of a series of S-functionalized 8-mercaptopguanine (8MG) analogues as substrate site inhibitors of HPPK and quantify binding against the *E. coli* and *S. aureus* enzymes (EcHPPK and SaHPPK). The results demonstrate that analogues incorporating acetophenone-based substituents have comparable affinities for both enzymes. Preferential binding of benzyl-substituted 8MG derivatives to SaHPPK was reconciled when a cryptic pocket unique to SaHPPK was revealed by X-ray crystallography. Differential chemical shift perturbation analysis confirmed this to be a common mode of binding for this series to SaHPPK. One compound (**41**) displayed binding affinities of 120 nM and 1.76 μ M for SaHPPK and EcHPPK, respectively, and represents a lead for the development of more potent and selective inhibitors of SaHPPK.



INTRODUCTION

With over 20 000 deaths a year from drug resistant infections in the U.S. alone, antibiotic resistance represents a growing worldwide healthcare concern.¹ A recent report by the World Health Organization² stressed the need for an increased global effort in combating antibiotic resistance, warning that a postantibiotic era, "... far from being an apocalyptic fantasy, is instead a very real possibility for the 21st century".² The inevitability³ of antibiotic resistance necessitates a constant influx of novel and effective antimicrobial agents. Compounded by a reduced financial incentive compared to other pharmaceuticals,^{4–6} regulatory changes have been made to streamline new antibiotic development and help replenish the R&D pipeline.^{4,6–8}

Folate is an essential vitamin for all living organisms. The reduced form, tetrahydrofolate, is a required cofactor in the synthesis of purines, pyrimidines, and amino acids.^{9,10} Humans depend on dietary folate and employ both passive and active mechanisms for uptake, while most microorganisms and lower eukaryotes, including plants, synthesize folate de novo. Several enzymes of the folate de novo pathway are without human homologues and therefore represent attractive targets¹¹ for

antibiotic development. For decades dihydropteroate synthase (DHPS) has been the target of the long-standing sulfonamide class of antibiotics, validating the pathway's utility for further antimicrobial development.¹² Sulfonamides and sulfones have been used to treat infectious diseases including malaria, *Toxoplasma gondii* encephalitis, *Pneumocystis jirovecii* pneumonia, tuberculosis, shigellosis, and *Staphylococcus aureus* infections, often coadministered with trimethoprim, an inhibitor of the folate-pathway enzyme dihydrofolate reductase (DHFR).^{13–18} Point mutations in the DHPS^{12,19} and DHFR genes^{20,21} have antagonized these combination therapies, motivating rational approaches toward newer, improved generations of DHFR²² and DHPS^{12,23,24} inhibitory lead compounds, potentially minimizing "sulfa allergy"²⁵ side effects. The increased spread of methicillin-resistant *S. aureus*²⁶ (MRSA) strains throughout hospitals and the community has been a major incentive in seeking new therapeutic approaches such as immunization strategies²⁷ and new pharmacological targets.^{28,29}

Received: January 3, 2016

Published: April 20, 2016

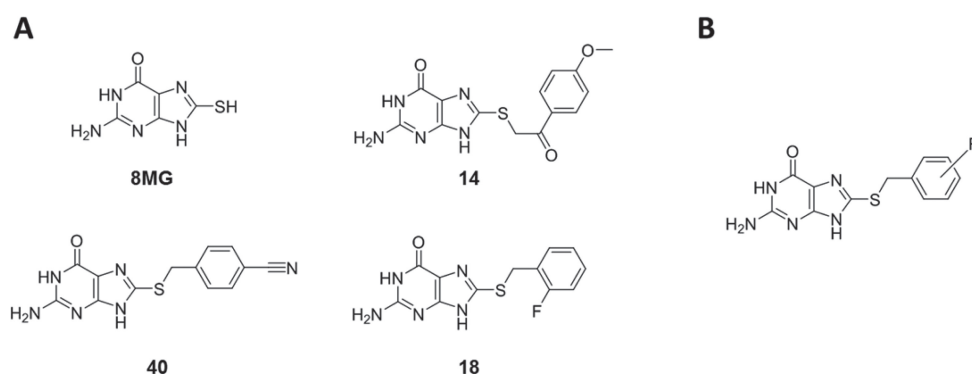


Figure 1. (A) Structures of selected HPPK inhibitors. (B) Scaffold of the HPPK inhibitors that were the main focus of this study.

Table 1. SAR for Alkyl-Substituted 8MG Derivatives^b

Compound	R	K_D SaHPPK (μ M)	K_D EcHPPK (μ M)
8MG	H	12.6 ± 1.3^a	160 ± 10
1		7.6 ± 0.2	89.2 ± 0.5
2		4.6 ± 0.3	68 ± 1
3		13.4 ± 0.2	82 ± 11
4		3.6 ± 1.1	97 ± 1
5		2.76 ± 0.01	52 ± 1
6		7.9 ± 1.6	23.3 ± 0.4
7		7.9 ± 1.1	n.d.
8		8.8 ± 1.0	50.0 ± 0.1
9		13 ± 1	59.4 ± 0.4
10		1.5 ± 0.1	53 ± 10
11		11 ± 1	74.2 ± 0.1
12		46 ± 4	120 ± 10

^aDennis et al.³³ ^bCompounds were tested in the presence of 1 mM ATP. Compounds 3, 4, and 12 were tested as racemic mixtures. n.d.: not determined.

Our research has focused on a structure-based approach to discover inhibitors of the enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) from *S. aureus*.^{30–34}

HPPK precedes DHPS in the folate-pathway, catalyzing the pyrophosphoryl transfer from the ATP cofactor to a 6-hydroxymethyl-7,8-dihydropterin (HMDP) substrate.³⁵ To our knowledge HPPK is currently not the target of any drugs and, like DHPS, has no human homologue, making it an attractive drug target.

HPPK is a monomeric³⁶ 18 kDa enzyme incorporating a three-layered α – β – α , thioredoxin-like fold. While the core of the protein remains fairly rigid during catalysis, three loop regions (loops 1–3) undergo conformational changes with major changes observed for loop 3 in particular.^{35,37} The ATP cofactor binds first³⁵ with loop 2 remaining partially mobile³⁸ and loop 3 extending away from the binding site.³⁵ Upon binding the pterin substrate, loop 3 closes over the binding site to create a complex hydrogen bond network formed between the three catalytic loops, rigidifying the ternary complex. This

seals the binding site, enabling the transfer of a pyrophosphate group from ATP to the substrate.^{35,38}

Despite the abundant structural and kinetic data available on HPPK catalysis^{39–41} within a range of species,^{31,42–46} few inhibitors have been reported. Substrate-site pterin analogues have been crystallized in complex with HPPK,^{43,47,48} and bitopic ligands targeting both the pterin and ATP cofactor pockets were reported by Shi et al.^{49–52} Our current research focuses on evolution of the 8-mercaptoguanine (8MG) core scaffold that we previously identified as a promising pterin-site inhibitor of *S. aureus* HPPK (SaHPPK) with good ligand efficiency ($K_D = 12.6 \mu\text{M}$ over 12 heavy atoms).^{31–33} Building on this initial discovery, we identified the commercially available acetophenone 8MG analogue, **14** (Figure 1), as a promising cofactor-dependent high affinity binder ($K_D = 1.1 \mu\text{M}$) via screening various 8MG-derived compounds.³³ Synthesis and testing of a focused library subsequently led to the identification of several other S8-functionalized derivatives of 8MG capable of binding SaHPPK, including some simple benzyl analogues such as **40** which exhibited significantly improved affinity ($K_D = 660 \text{ nM}$).³³

This body of work presents the further development of 8MG-based inhibitors through synthesis and structural studies, with a focus on a series of benzyl-substituted derivatives. An unexpected binding pocket, specific to the SaHPPK structure, was discovered and probed, revealing a means to further improve affinity as well as rationalize the selectivity observed for the SaHPPK enzyme with this class of compound.

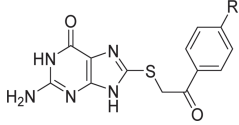
RESULTS AND DISCUSSION

Surface plasmon resonance (SPR) analysis was used to determine the equilibrium binding constant (K_D) for the interaction of compounds with the *E. coli* and *S. aureus* enzymes, employing biotinylated-HPPK captured to a streptavidin-coated chip. As observed previously,^{33,53} initial screening confirmed that compounds generally bound with significantly higher affinity to cofactor-bound HPPK, compared to the *apo* enzyme (data not shown). HPPK is saturated with ATP in vivo ($K_D \approx 40 \mu\text{M}$ ^{31,41}), and therefore SPR screening was performed exclusively in the presence of saturating levels of cofactor (1 mM ATP). Compounds were tested in a dose-dependent manner with the highest concentration ranging between 3 and 300 μM to cover both sub and supra K_D concentrations. Where solubility or affinity was limiting, compounds were globally fitted using a known compound that was approaching saturation to the protein.

Owing to differences in instrumentation and experimental setup, including the method of surface coupling and enzyme capture, we measured the affinity of some previously reported 8MG derivatives^{33,53} alongside our new series of derivatives for direct reference (Tables 1, 2, and 3). Notably, the cysteine-conjugated biotinylated SaHPPK used in this study returned affinities approximately 2-fold tighter than those previously obtained with SaHPPK immobilized directly via one or more lysine residues.³³

S-Alkylation of 8MG Yields Marginal Improvements in Binding Affinity. In prior HPPK work,^{32,33,53} testing of S-functionalized 8MG-derivatives with nonaromatic substituents was limited to four compounds. As part of the current study, an expanded series of S8-alkylated 8MG-derivatives was synthesized to better probe the HPPK binding site. SPR testing of these compounds (Table 1) revealed significantly weaker binding to both SaHPPK and *E. coli* HPPK (EcHPPK)

Table 2. Binding of Acetophenone-Substituted 8MG Derivatives^a



compd	R	K_D (μM)	
		SaHPPK	EcHPPK
13	Br	0.57 ± 0.01	0.70 ± 0.01
14	OCH_3	0.60 ± 0.04	0.82 ± 0.04
15	phenyl	0.69 ± 0.11	0.92 ± 0.01
16	H	1.2 ± 0.1	2.0 ± 0.1

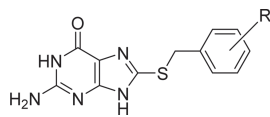
^aCompounds were tested in the presence of 1 mM ATP. The compounds **14**,^{33,53} **15**,³³ and **16**^{33,53} are previously discovered.

compared to the aromatic derivatives (vide infra), and only marginal improvements in affinity relative to the parent compound, 8MG ($K_D = 12.6 \mu\text{M}$ ³³ and 160 μM in SaHPPK and EcHPPK, respectively).

The series included compounds with hydroxyl, ether, and amide groups within the alkyl substituents. The alkyl alcohols, **1**, **2**, and **4**, bound SaHPPK more tightly than did 8MG ($12.6 \mu\text{M}$ ³³), with the best analogue, **4**, possessing a K_D of 3.6 μM . Compounds with alkyl ether moieties (**5**, **6**) were also well tolerated, with **5** yielding a K_D of 2.76 μM . Compounds with amide moieties were reasonably tolerated, with the pyrrolidinone **7** giving a K_D of 7.9 μM . Interestingly, the carboxamide analogues gave the sharpest SAR. The secondary carboxamides **8** and **9** demonstrated K_D (SaHPPK) values of 8.8 and 13 μM , respectively, with a selectivity index (SI) of ~ 5 for SaHPPK vs EcHPPK binding, while the tertiary carboxamides **10** and **11** gave K_D (SaHPPK) values of 1.5 and 11 μM , respectively. Introducing branching at the α -position of the carboxamide through a methylene group led to a significant decrease in affinity, with **12** returning a K_D of 46 μM . In terms of activity against EcHPPK, it was striking that all compounds were significantly less active. In this respect, the tertiary amide **10** represents an interesting new lead for selective inhibitors of SaHPPK due to its high ligand efficiency ($\Delta G/\text{number of heavy atoms}$ is $0.43 \text{ kcal}\cdot\text{mol}^{-1}\cdot(\text{heavy atom})^{-1}$) and its K_D (EcHPPK) of 53 μM , providing an SI of ~ 35 for SaHPPK binding.

8MG-Acetophenone Derivatives Bind SaHPPK and EcHPPK with Comparable Affinity and with a Similar Binding Pose. In contrast to the preferential SaHPPK binding observed for the alkyl-substituted compounds, SPR testing of compounds possessing an acetophenone substitution from the S8-position (**13–16**) revealed comparable SaHPPK and EcHPPK binding affinities (Table 2). In particular, compounds **13–15** were noted to bind to SaHPPK ($K_D = 0.57, 0.60, 0.69 \mu\text{M}$, respectively) and EcHPPK ($K_D = 0.70, 0.82, 0.92 \mu\text{M}$, respectively) with remarkably similar affinities. The apparent 13-fold weaker EcHPPK vs SaHPPK binding reported previously for **14**^{33,53} is attributable instead to the different SPR conditions employed in the two separate studies. Specifically, a subsaturation concentration of the cofactor analogue AMPCPP was employed in the EcHPPK SPR experiments of Yun et al.⁵³ compared to saturated levels of ATP in our studies.³³

The comparable SaHPPK and EcHPPK binding affinities of the acetophenone series reflect the similar pose adopted by

Table 3. SAR for Benzyl-Substituted 8MG Derivatives^a

compd	K_D (μ M)		ortho	meta	para
	SaHPPK	EcHPPK			
17	0.46 \pm 0.03	6.9 \pm 0.4	H	H	H
18	0.18 \pm 0.03	1.5 \pm 0.2	F	H	H
19	1.3 \pm 0.1	nd	Cl	H	H
20	1.5 \pm 0.1	8.5 \pm 1.2	Br	H	H
21	0.47 \pm 0.08	2.5 \pm 0.4	Me	H	H
22	0.94 \pm 0.24	7.1 \pm 0.1	CF ₃	H	H
23	0.54 \pm 0.05	5.7 \pm 0.8	NO ₂	H	H
24	0.51 \pm 0.01	23 \pm 2	CN	H	H
25	0.22 \pm 0.05	2.6 \pm 0.3	H	F	H
26	0.57 \pm 0.01	1.7 \pm 0.1	H	Cl	H
27	3.1 \pm 0.3	7.1 \pm 0.2	H	Br	H
28	0.38 \pm 0.01	1.7 \pm 0.2	H	Me	H
29	0.65 \pm 0.07	1.8 \pm 0.1	H	OMe	H
30	1.61 \pm 0.01	5.7 \pm 0.1	H	OCF ₃	H
31	1.4 \pm 0.3	5.8 \pm 0.5	H	NO ₂	H
32	0.77 \pm 0.13	3.5 \pm 0.5	H	CN	H
33	2.6 \pm 0.1	26 \pm 2	H	COOH	H
34	0.30 \pm 0.02	3.9 \pm 0.3	H	H	F
35	1.24 \pm 0.04	4.3 \pm 0.2	H	H	Cl
36	0.95 \pm 0.10	2.13 \pm 0.03	H	H	Br
37	0.52 \pm 0.03	2.2 \pm 0.2	H	H	OMe
38	1.93 \pm 0.04	nd	H	H	CF ₃
39	2.1 \pm 0.4	5.2 \pm 0.2	H	H	OCF ₃
40	0.33 \pm 0.03	4.8 \pm 0.3	H	H	CN
41	0.12 \pm 0.02	1.76 \pm 0.05	F,F	H	H
42	2.3 \pm 0.1	6.9 \pm 0.6	H	Me,Me	H
43	1.2 \pm 0.1	nd	H	CF ₃ ,CF ₃	H
44	0.841 \pm 0.001	6.1 \pm 0.1	Me	Me (3-position)	H
45	1.7 \pm 0.2	12.6 \pm 0.3	Me	Me (5-position)	H
46	2.8 \pm 0.1	10.7 \pm 0.4	F	Me (3-position)	H
47	0.21 \pm 0.02	0.965 \pm 0.003	F	Me (5-position)	H
48	0.26 \pm 0.05	1.4 \pm 0.1	F	H	F
49	0.30 \pm 0.01	0.95 \pm 0.11	F	H	OMe
50	0.59 \pm 0.10	2.7 \pm 0.3	F	H	CN
51	1.97 \pm 0.01	16.3 \pm 0.3	Me	H	F
52	1.2 \pm 0.1	38 \pm 3	CN	H	F
53	0.29 \pm 0.05	2.0 \pm 0.0	H	F	F
54	3.6 \pm 0.2	10.1 \pm 0.2	H	Me	F
55	2.1 \pm 0.3	5.3 \pm 0.9	H	CN	F

^aCompounds were tested in the presence of 1 mM ATP. The compounds 17,³³ 18,³³ 22,⁵³ and 40^{33,53} are previously discovered. nd: not determined.

these ligands upon binding to the two enzymes^{33,53} as exemplified by the crystal structures of **13** in complex with SaHPPK and EcHPPK (Figure 2). In both structures, the thiol extension projects out from the substrate pocket past loop 2, making interactions with R121, F123, and the two loop 2 residues V46 and G47 (L45 and G46 in EcHPPK), with additional interactions to the loop 3 residue W89, in EcHPPK (Figure 2). Overall, these results strongly indicate that selectivity gains are not anticipated through simple extension of 8MG with acetophenone-type pendent groups.

S-Benzoylation of 8MG Induces a Novel Mode of Binding to SaHPPK. It was previously shown that addition of benzyl substituents to the S8 position of 8MG results in

stronger SaHPPK binding than acetophenone substitution³³ but curiously 2-fold *weaker* binding to EcHPPK.⁵³ Multiple crystallization attempts failed to deliver suitable crystals of our best benzyl lead (compound **18**) in complex with SaHPPK for high resolution structural evaluation. This prompted us to explore the synthesis and binding properties of further S-benzyl analogues.

In addition to several monosubstituted benzyl bromides, various disubstituted benzyl bromides were used to alkylate 8MG (Figure 1B) to determine whether we could observe additive SAR (Table 3). Our compound set included compounds with halogen and electron withdrawing and electron donating groups ring substituents (ortho, meta, and

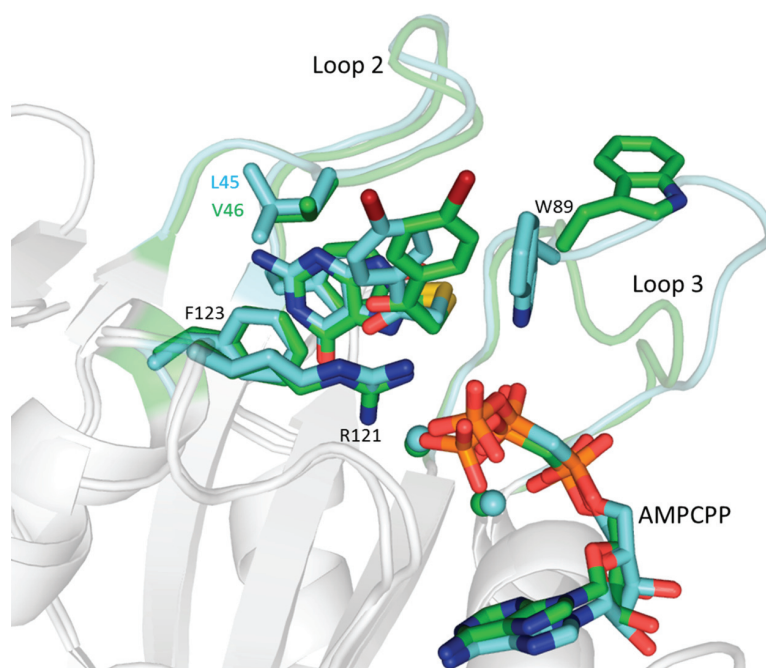


Figure 2. X-ray structures of **13** and AMPCPP complexed to EcHPPK (cyan) and SaHPPK (green).

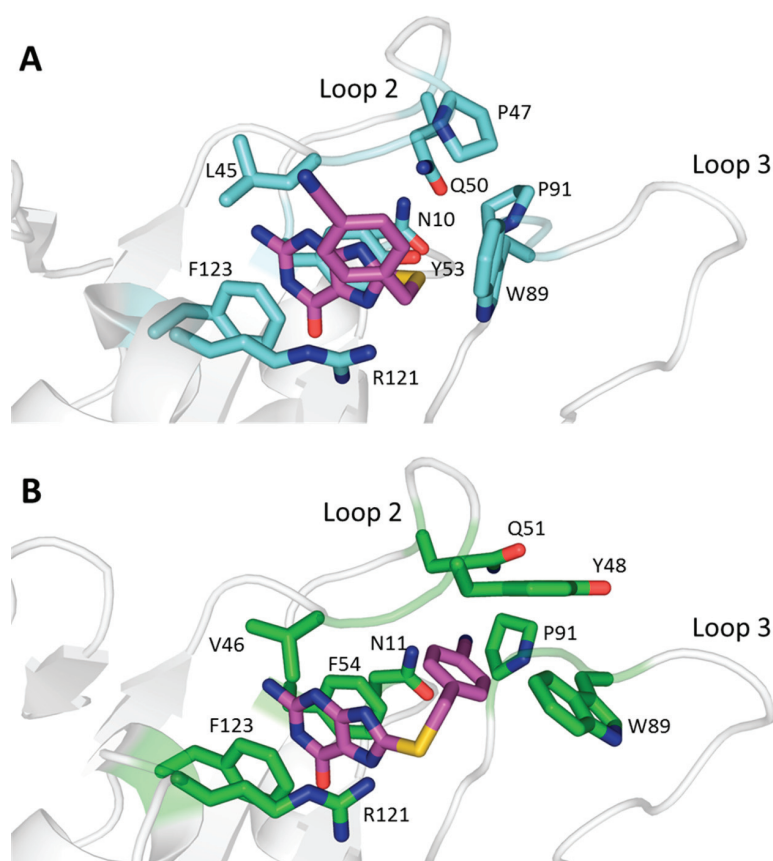


Figure 3. X-ray structures of **40** and AMPCPP bound to (A) EcHPPK⁵³ or (B) SaHPPK. Selected key residues for each protein are highlighted in cyan and green, respectively. **40** is displayed in purple. AMPCPP and coordinated ions are omitted for clarity.

para positions). As we investigated this focused benzyl 8MG library, it became increasingly difficult to reconcile the

emerging SAR based on the ligand binding poses observed in earlier structural studies.^{33,53}

Efforts to acquire a structural explanation for the species dependent SAR included an additive crystallization screen, performed with various single- and disubstituted benzyl 8MG derivatives in the presence of AMPCPP and SaHPPK. A new crystal form with AMPCPP and *p*-cyanobenzyl-substituted 8MG **40** bound diffracted to 1.96 Å resolution, showing two protomers in the asymmetric unit (space group $P2_1$). Within this structure, the guanine moiety of **40** binds in a similar fashion to that observed in prior 8MG and 8MG-derivative complexes,^{31–33,53} making a total of six hydrogen bonds with the protein and π -stacking between the aromatic rings of F54 and F123. Notably, however, the thiol extension displays a novel orientation compared to the acetophenone-bearing derivative **13** (Figure 2), with residues 47–51 and 84–91 of loops 2 and 3, respectively, displaced by up to 17 Å from their position in the **13**/AMPCPP/SaHPPK complex, enlarging a small pocket to accommodate the benzyl extension (Figure 3B). The sulfur atom in **40** is translated ~ 1 Å further away from the pocket relative to that in **13** to allow the extension to fit. Outside these regions, the protein and cofactor are relatively unchanged, including the loop 3 catalytic arginine residues R83 and R92 (rmsd of 0.6 and 0.7 Å, respectively, relative to the **13**/AMPCPP/SaHPPK complex). The benzyl ring is involved in hydrophobic interactions with the side chains of residues P45, Y48, Q51, F54, and W89, while the nitrile group interacts with residues G90 and P91, and the side chains of N11 and Q51.

The structure contrasts with the binding mode observed in the previously reported **40**/AMPCPP/EcHPPK complex⁵³ (Figure 3A), in which the thiol extension is positioned as per the acetophenone derivatives, interacting with residues R121, F123, L45, G46, and W89, while the nitrile is solvent-exposed and forms no significant interactions. Eight additional SaHPPK/EcHPPK crystal structures (Table 4, Figures 4 and 5) further confirm that the formation of this new binding pocket is only evident upon binding of S-benzyl substituted 8MG derivatives to SaHPPK.

Table 4. Summary of Crystallographic Data for Ligand-Bound AMPCPP/HPPK Ternary Complexes

compd	protein	space group	resolution (Å)	$R_{\text{work}}/R_{\text{free}}$	PDB accession code
13	SaHPPK	$P6_1$	1.72	18.1/20.7	SETV
	EcHPPK	$P2_1$	1.05	10.7/12.5	SETP
18	EcHPPK	$P2_1$	1.09	11.6/13.6	SETK
24	EcHPPK	$P2_1$	1.82	20.2/24.1	SETL
34	SaHPPK	$P2_1$	1.32	12.4/15.2	SETR
	EcHPPK	$P2_1$	1.46	20.4/23.9	SETM
35	SaHPPK	$P2_1$	1.95	18.9/21.8	SETS
	EcHPPK	$P2_1$	1.40	18.2/21.9	SETN
40	SaHPPK	$P2_1$	1.96	18.2/21.6	SETQ
50	SaHPPK	$P2_1$	1.55	15.8/19.2	SETT
	EcHPPK	$P2_1$	1.07	12.9/15.0	SETO

Binding affinities of the benzylated compounds for SaHPPK and EcHPPK are displayed in Table 3. The unsubstituted scaffold, **17**, gave K_D values of 460 ± 30 nM and 6.9 ± 0.4 μ M for SaHPPK and EcHPPK, respectively. This 15-fold difference in affinity can now be rationalized in terms of the new crystallographic data: in EcHPPK, the lack of a carbonyl group in the linker appears to remove the favorable π interactions to the guanidinium side chain of the conserved⁴⁷ R121 residue, resulting in a ~ 3 -fold loss of affinity compared with **16** ($K_D =$

2.0 μ M). In SaHPPK, however, the shortened linker is accommodated by the new pocket, resulting in a ~ 3 -fold improvement in affinity compared with **16** ($K_D = 1.2$ μ M). Compounds with benzylic extensions were all found to be weaker EcHPPK binders due to an inability to access a binding pocket akin to that observed in SaHPPK, with the highest affinity belonging to **49** ($K_D = 0.95$ μ M). Conversely, submicromolar SaHPPK affinities were consistently observed for this series of compounds.

SAR Analysis for EcHPPK Binding. For EcHPPK binding, only small ortho-substituents are favored, with the fluoro (**18**, $K_D = 1.5$ μ M) and methyl (**21**, $K_D = 2.5$ μ M) groups improving affinity relative to the unsubstituted benzylated compound (**17**, $K_D = 6.9$ μ M). The slightly larger bromo (**20**, $K_D = 8.5$ μ M), trifluoromethyl (**22**, $K_D = 7.1$ μ M), and nitro (**23**, $K_D = 5.7$ μ M) groups maintain affinity, while a loss of affinity occurs with a nitrile substituent at the ortho-position (**24**, $K_D = 23$ μ M). This was also apparent for our disubstituted analogue **52**, where introduction of a nitrile at the ortho position resulted in a significant decrease in affinity (~ 10 -fold) relative to the equivalent monosubstituted compound **34**.

The crystal structures of **18**/AMPCPP/EcHPPK and **24**/AMPCPP/EcHPPK reveal the reason for the improved affinity accompanying *o*-fluorine substitution but loss upon *o*-nitrile substitution (Figure 4). In previously reported EcHPPK ternary complexes of analogous compounds,⁵³ the conserved residue W89 π -stacked against the aromatic moiety of the thiol extension. This is similar to what is observed in the ternary complexes of AMPCPP/EcHPPK with **13**, **18**, **34**, **35**, and **50**, presented in this work. In the **18**/AMPCPP/EcHPPK complex, the *o*-fluoro group favorably interacts with the side chain of residues R121 and F123 (Figure 4A), leading to improved affinity relative to the unsubstituted form. In the ternary complex of **24** (Figure 4B), however, the larger nitrile group is unable to fit in this orientation, resulting in a flipping of the *o*-substituent to the other side of the ring so that it is oriented away from residues R121/F123. This repositioning precludes the ring stacking interaction with W89, resulting in increased apparent mobility and a corresponding loss of affinity. In three of the four protomers within the asymmetric unit, this mobility manifests as a lack of electron density for W89 and adjacent residues, with density in the remaining protomer being weak and discontinuous for this region. The nitrile group also displaces P47, resulting in loop 2 backbone and side chain movements in excess of 4 and 13 Å, respectively, relative to the positions adopted in the **18**/AMPCPP/EcHPPK complex.

Meta-substitution is better tolerated by EcHPPK, with an improvement in affinity demonstrated for compounds with fluoro (**25**, $K_D = 2.6$ μ M), chloro (**26**, $K_D = 1.7$ μ M), methyl (**28**, $K_D = 1.7$ μ M), methoxy (**29**, $K_D = 1.8$ μ M), and nitrile (**32**, $K_D = 3.5$ μ M) substituents. Affinity is maintained with a bromo (**27**, $K_D = 7.1$ μ M), trifluoromethoxy (**30**, $K_D = 5.7$ μ M), and nitro (**31**, $K_D = 5.8$ μ M) substituent. A significant loss of affinity was observed for the meta-carboxylic acid substituent (**33**, $K_D = 26$ μ M). This is believed to be a result of the increased size and charged nature of the carboxylic acid moiety, which may flip the benzyl ring and perturb interactions with W89 in a manner similar to that described above.

Para-substitution is observed to have a negligible effect on binding compared with ortho- or meta-substitution, likely a result of the more solvent-exposed nature of this position. A *p*-bromo (**36**) or *p*-methoxy (**37**) substituent improved affinity 3-fold ($K_D = 2.13$ and 2.2 μ M, respectively) over the parent

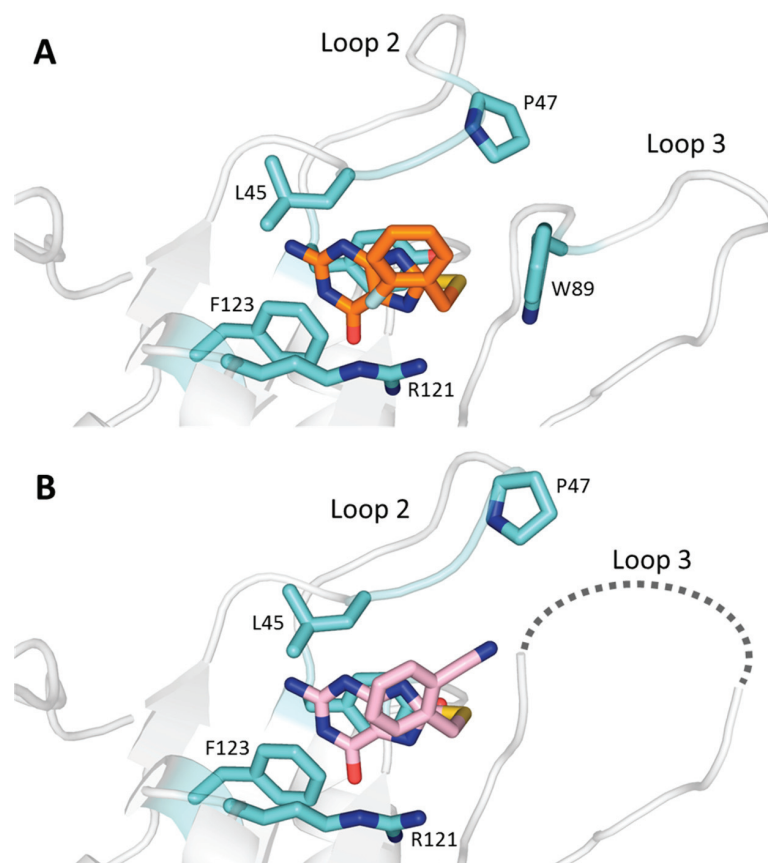


Figure 4. X-ray structures of EcHPPK bound to AMPCPP and (A) **18** or (B) **24**. The loop 3 residues that are absent in the electron density (86–90) of the **24**/AMPCPP/EcHPPK complex are represented by a dashed gray line. AMPCPP and coordinated ions are omitted for clarity.

compound, while the other para-substituted derivatives synthesized exhibited moderate K_D values of 3.9–5.2 μM .

Halogen groups proved most beneficial for EcHPPK binding, with fluorine, chlorine, and bromine returning the highest affinities for substituents at the ortho, meta, and para positions, respectively. The highest EcHPPK affinities were observed for disubstituted derivatives possessing an *o*-fluoro group and either an *m*-methyl (**47**, $K_D = 0.965 \mu\text{M}$), *p*-fluoro (**48**, $K_D = 1.4 \mu\text{M}$), or *p*-methoxy (**49**, $K_D = 0.95 \mu\text{M}$) substituent.

SAR Analysis for SaHPPK Binding. In SaHPPK, the new binding pocket for the thiol extension is quite small and sterically restricted, so only smaller substituents are favored or tolerated at each position. At the ortho position, a fluorine atom is the only substituent to improve binding affinity over the unsubstituted compound **17** ($K_D = 0.46 \mu\text{M}$), as per the previously reported monosubstituted compound **18** ($K_D = 0.18 \mu\text{M}$).³³ The crystal structure of the **50**/AMPCPP/SaHPPK complex indicates that this may be the result of a favorable interaction with the catalytic R92 residue (Figure 5A). Ortho-substitution with methyl (**21**, $K_D = 0.47 \mu\text{M}$), nitro (**23**, $K_D = 0.54 \mu\text{M}$), or nitrile (**24**, $K_D = 0.51 \mu\text{M}$) groups maintained affinity, while chloro (**19**, $K_D = 1.3 \mu\text{M}$), bromo (**20**, $K_D = 1.5 \mu\text{M}$), and trifluoromethyl (**22**, $K_D = 0.94 \mu\text{M}$) substituents caused an approximately 2- to 3-fold loss in affinity. At the meta-position, fluoro (**25**, $K_D = 0.22 \mu\text{M}$) and methyl (**28**, $K_D = 0.38 \mu\text{M}$) groups improved affinity, while chloro (**26**, $K_D = 0.57 \mu\text{M}$), methoxy (**29**, $K_D = 0.65 \mu\text{M}$), and nitrile (**32**, $K_D = 0.77 \mu\text{M}$) groups maintained affinity. In contrast, the larger groups (bromo (**27**, $K_D = 3.1 \mu\text{M}$), trifluoromethoxy (**30**, $K_D =$

1.61 μM), nitro (**31**, $K_D = 1.4 \mu\text{M}$), and carboxylic acid (**33**, $K_D = 2.6 \mu\text{M}$)) decreased affinity for SaHPPK. At the para position, a nitrile substituent slightly improved affinity (**40**, $K_D = 0.33 \mu\text{M}$), which may be a result of the aforementioned interactions with the side chains of N11 and Q51 (Figure 3B). The *p*-fluoro derivative, **34**, also displayed improved binding ($K_D = 0.30 \mu\text{M}$), while *p*-chloro (**35**, $K_D = 1.24 \mu\text{M}$) or *p*-bromo (**36**, $K_D = 0.95 \mu\text{M}$) substitution reduced binding by approximately 2- to 3-fold, possibly due to their larger atomic radii. The structures of **34** and **35** complexed with AMPCPP/SaHPPK are similar (Figure 5B), the main difference being in the positioning of the phenyl ring to accommodate the larger sized chlorine atom. Other derivatives tested included one with a *p*-methoxy substituent (**37**, $K_D = 0.52 \mu\text{M}$) which maintained the activity of the unsubstituted parent compound, while a larger *p*-trifluoromethoxy group (**39**, $K_D = 2.1 \mu\text{M}$) reduced affinity.

To determine whether we could improve affinity by introducing disubstituted benzyl moieties, we synthesized a focused library of various disubstituted analogues. Of the disubstituted compounds, those possessing *o*-fluoro substitution with additional meta- or para-substituents generally showed the highest SaHPPK affinities (e.g., compound **47**, possessing *o*-F, *m*-methyl substitution, $K_D = 0.21 \mu\text{M}$). However, the only compound that demonstrated improved binding relative to **18** was the 2,5-difluoro derivative **41** ($K_D = 120 \text{ nM}$). Despite their somewhat lower affinity for SaHPPK, the derivatives possessing *o*-nitrile substitution (**24**, **52**) possessed the greatest SIs (**45** and **32**, respectively) of the

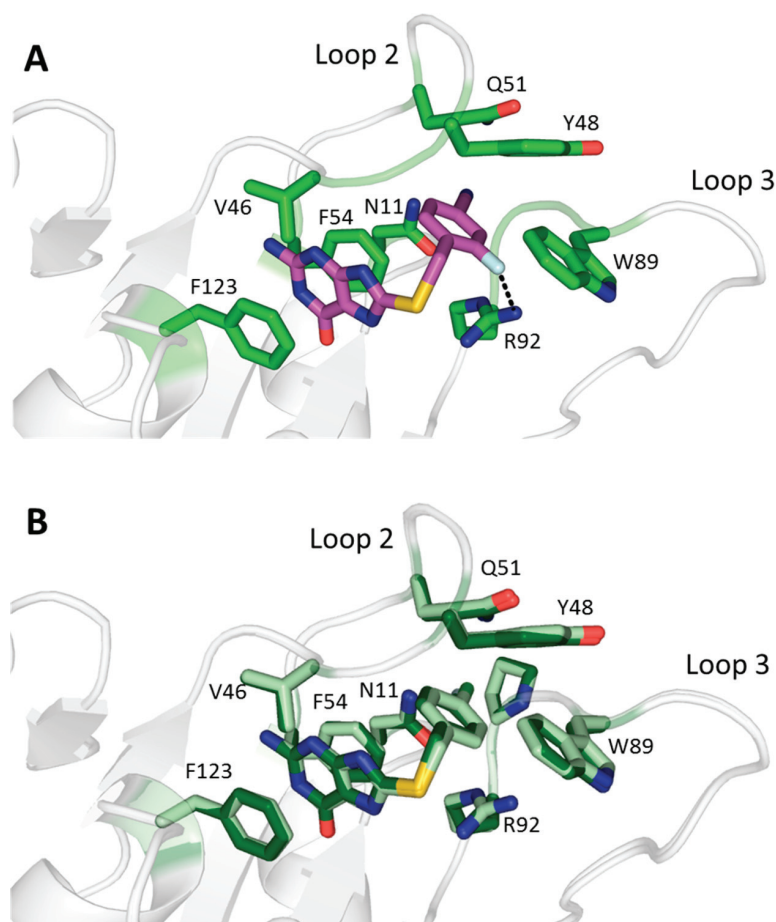


Figure 5. (A) X-ray structure of **50**/AMPCPP/SaHPPK. Interaction to the side chain of R92 is represented as a dashed black line. (B) X-ray structures of AMPCPP/SaHPPK complexed to **34** (dark green) or **35** (light green). AMPCPP and coordinated ions are omitted for clarity.

benzyl series and may therefore represent robust leads in the pursuit of species-selective HPPK inhibitors.

Sequence Differences Provide a Rationale for Differences in SaHPPK and EcHPPK Binding. The SAR analysis in this body of work lays a foundation for producing compounds with increased selectivity to SaHPPK. The unique binding pocket observed in the crystal structures of SaHPPK with the benzylated 8MG derivatives bound explains the divergent SAR for SaHPPK vs EcHPPK binding. The two key interacting residues are Y53 (F54 in SaHPPK) and P47 (Y48 in SaHPPK); the additional hydroxyl group found in Y53 appears to be incompatible with the benzyl moiety due to steric interaction, while the proline residue, P47, may limit the conformational freedom of loop 2 (see Figure 3). The dihedral angles observed for the equivalent SaHPPK residue, Y48 ($\varphi = -156.9^\circ$, $\psi = -164.1^\circ$), are prohibited in proline, and the loop 2 rearrangements that form the cryptic pocket in SaHPPK are unlikely to be energetically feasible. Another point of difference is that loop 3 is longer in EcHPPK, with the additional residue, A86, potentially hindering the formation of the binding pocket observed in SaHPPK.

NMR Analysis of SaHPPK Complexes. Cococrystallization of various ligands with AMPCPP and EcHPPK was generally successful for most compounds tested. In comparison, cococrystallization of the same ligands with AMPCPP and SaHPPK was only successful for a subset of compounds tested.

Only the para-substituted benzylated 8MG derivatives resulted in crystals of sufficiently high diffraction quality, specifically those possessing either a fluoro (**34**), chloro (**35**), or nitrile (**40**, **50**) group. With the parent analogue (unsubstituted benzyl derivative **17**) and compounds possessing mono ortho- or meta-substitution, we were unable to obtain any appreciable crystal growth.

To investigate ligand binding to SaHPPK for compounds lacking crystallographic data, the ^1H N, ^{15}N , and ^{13}C A backbone resonances for the **17**/AMPCPP/SaHPPK and **41**/AMPCPP/SaHPPK complexes, in addition to the crystallized **40**/AMPCPP/SaHPPK and **50**/AMPCPP/SaHPPK complexes, were assigned using a HNCA triple-resonance 3D NMR experiment. Binding of these four compounds to the AMPCPP-loaded SaHPPK gave some large chemical shift changes in the ^{15}N HSQC spectra in comparison to the spectrum of the AMPCPP/8MG parent complex (Figure 6). Binding was observed to be in the slow exchange regime, consistent with the measured affinities (Table 3). Comparison of these data with the previously assigned ternary complexes of **14** and **18**³³ provides firm evidence that the ligand binding mode observed in the crystal structures of SaHPPK with the benzylated compounds bound is maintained in the absence of para-substitution.

To probe the position of the thiol extension specifically, we compared the ^1H – ^{15}N chemical shift perturbations (CSP)

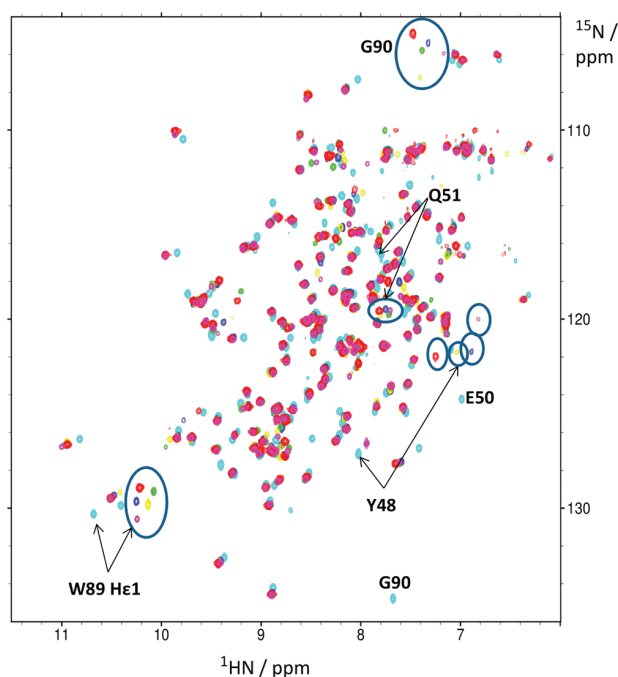


Figure 6. Superposition of various 2D ^{15}N FAST-HSQC spectra of $\sim 100\ \mu\text{M}$ SaHPPK samples in complex with AMPCPP. The spectrum of SaHPPK in complex with 8MG/AMPCPP (cyan) is compared with those from five 8MG benzyl derivatives: *o*F/AMPCPP (blue), *p*CN/AMPCPP (green), *o*F-*p*CN/AMPCPP (red), and di-*o*F/AMPCPP (yellow). Several 8MG/AMPCPP assignments are identified (cyan) and the corresponding clustering of peaks in the spectra of the benzyl derivatives highlighted with a ring. Spectra were recorded at a field strength of 14.1 T and at 22 $^{\circ}\text{C}$.

observed in the ^{15}N HSQC spectra of each of the four benzylated derivatives in complex with AMPCPP/SaHPPK (Figure 6) to those observed in the 2D NMR spectrum of the bound guanine core in 8MG/AMPCPP/SaHPPK. All four compounds share the same CSP signature (Figure 7A), indicative of a similar pose for the thiol extension. In comparison, the signature for the thiol extension in 14/AMPCPP/SaHPPK (shown in red in Figure 7A) is markedly different. Large CSPs are noted for amides Y48, T49, and Q51 in loop 2 and G90 in loop 3 (Figure 7A) in the spectra of the benzylated compounds in complex with AMPCPP/SaHPPK, as compared to the CSPs for the 14/AMPCPP/SaHPPK spectra. These residues map to the cryptic binding pocket within the 50/AMPCPP/SaHPPK structure (Figure 7B), confirming a similar pose for this series of compounds. Previously we had interpreted the differential CSPs from 2D NMR spectra of compound 18/AMPCPP/SaHPPK³³ as arising from backbone conformational changes in loop 2, in view of a docked model to the acetophenone 14/AMPCPP/SaHPPK structure (PDB code 4CRJ). Notably, in this structure, the pocket is too small to accommodate even the simple benzyl extension.

Compared with the 8MG/AMPCPP/SaHPPK case, several of the amide peaks in loop 2 were severely broadened or not observed (Figure 7C). In all spectra acquired of the benzyl series in complex with AMPCPP/SaHPPK, G47 was not observed, and while it was observed for 8MG, it was partially broadened in the 14/AMPCPP/SaHPPK spectrum. The observation of broadening indicates local μs – ms mobility of the ligand, the G47 backbone amide, or from a proximal group.

Qualitatively, it appears that the nitrile group may stiffen the loop more so than the other substituents investigated. We suggest that these active site loop dynamics hinder SaHPPK crystallization and that the para-substituent's interactions with residues of all three main loop regions (i.e., N11, Q51, G90, and P91) could promote crystallization. This loop stabilization is not necessarily thermodynamically favorable for ligand binding, however, and in the case of 35 (a compound successfully crystallized with SaHPPK), *p*-chloro substitution was observed to produce a ~ 3 -fold loss in affinity compared to the equivalent hydro compound.

Molecular Docking of HPPK Inhibitors. The suitability of molecular docking to determine ligand-bound structures was assessed for the EcHPPK and SaHPPK enzymes. Native ligand docking was performed for each X-ray determined complex using the docking program Glide⁵⁴ in both “standard precision” (SP) and “extra precision” (XP) modes. For SaHPPK complexes, both modes displayed a similar accuracy. All 8MG-benzylated poses gave a heavy atom root-mean-square deviation (RMSD) of $\leq 0.26\ \text{\AA}$ to their crystallographically determined position, while the acetophenone, 13, docked incorrectly ($\text{rmsd} > 2.0\ \text{\AA}$) in both modes (Table S3). For EcHPPK, XP mode performed noticeably better than SP mode for 5 of the 6 ligands (Table S4).

Cross-docking of ligands was then performed into the structures that were of the highest resolution, namely, the 34-bound SaHPPK (PDB code SETR) and the 13-bound EcHPPK (PDB code SETP), using XP mode. In EcHPPK only compounds 34 and 35 bound accurately during cross-docking. The docked model of 18, for example, asserts a flipped orientation of the ring compared to the crystal structure, the *o*-fluorine interacting with W89 (Figure 8A). Therefore, to account for the changes in the EcHPPK active site structure in response to different ligands, notably around W89, the X-ray structure with the most similar derivative was chosen for the docking. For example, the X-ray structure of the ortho-fluorinated 18 was used when docking the di-ortho-fluoro 41 into EcHPPK. In this docked model (Figure 8B), the phenyl ring of 41 is shifted $\sim 0.5\ \text{\AA}$ from 18 to accommodate the additional fluorine atom, which makes interactions to G46 and W89.

In contrast to EcHPPK, cross-docking was successful in SaHPPK for all 8MG-benzyl derivatives (rmsd values of $\leq 0.40\ \text{\AA}$) and suggests that molecular docking may be suitable to investigate the binding of similar compounds that lack structural data for SaHPPK. Compounds that lacked crystal structures, including the lead compound 41, were therefore docked into SaHPPK by this method to examine the interactions of the benzene substituents. 41 docks to SaHPPK with a binding mode analogous to that observed in the crystal complex of 50 in which the fluorine atom interacts with W89 and R92, with the additional fluorine atom interacting with P45, V46, and position N9 of the guanine ring (Figure 8C). When docked into SaHPPK, benzylic derivatives with meta-substitution rotated the benzyl ring such that the substituent is oriented as per the para-substituted compounds (Figure 8D). If meta-substituted compounds do indeed bind in this manner, the docking helps to rationalize why no additive affinity was observed in the *m,p*-difluoro analogue, 53, over the monosubstituted compounds 25 and 34.

These molecular modeling findings reiterate the sterically restricted nature of the new pocket in SaHPPK and further clarify the observation that functionalization from the phenyl

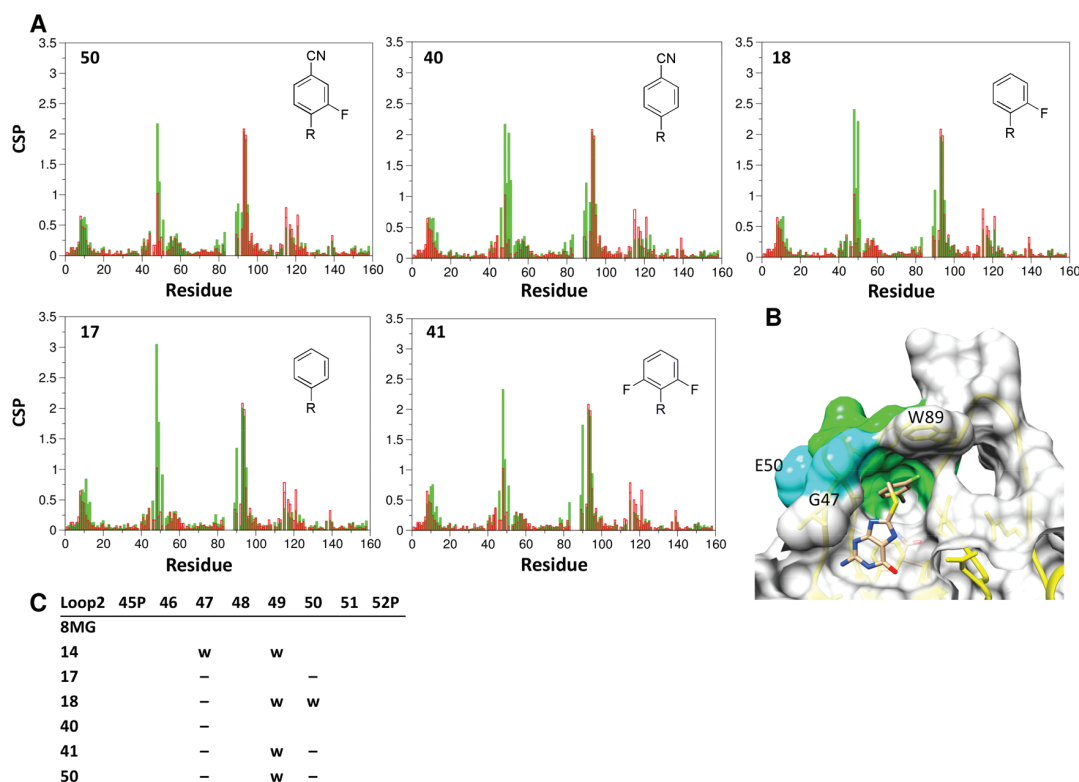


Figure 7. (A) Weighted CSPs between 2D ^{15}N HSQC spectra of SaHPPK in complex with various benzyl substituents/AMPCPP compared to the spectrum of SaHPPK in complex with 8MG/AMPCPP (green), where R is 8MG. For comparison the CSPs between 14/AMPCPP and 8MG/AMPCPP spectra are also shown (red). (B) Surface representation of part of the X-ray structure of 50/AMPCPP showing the cryptic pocket. The large CSPs for the simple benzyl (17, bottom left panel) compared to 14 have been mapped in green. The peaks for E50 and G47 are not observed in the 2D NMR spectra of 17/AMPCPP, and these residues are colored cyan. (C) 2D ^{15}N HSQC peaks from loop 2 residues which displayed reduced intensity (w) or were not observed (–) relative to the 8MG/AMPCPP 2D spectrum.

ring is limited to small, hydrophobic substituents. This suggests that future compounds may benefit instead from the use of phenyl ring isosteres to modulate and enhance interactions rather than investigating further substituents.

CONCLUSION

Exploiting the less conserved and often plastic or “cryptic” regions surrounding active sites is a worthwhile approach to enhance affinity and selectivity of an inhibitor compared with targeting the more rigid, conserved residues that provide vital recognition sites for the endogenous ligand. Structural data for a series of novel S-benzylated 8MG derivatives have revealed a distinct SaHPPK binding mode compared to that observed for EcHPPK, which has allowed us to rationalize previously conflicting and puzzling SAR. While only a relatively modest 4-fold improvement in affinity for SaHPPK ($K_D = 120$ nM) over previously reported 8MG-derivatives³³ has been achieved in this study, compound 41 has admirable ligand efficiency ($0.44 \text{ kcal}\cdot\text{mol}^{-1}\cdot(\text{heavy atom})^{-1}$) and is also a structurally well characterized scaffold to aid progression. The linker region provides a promising site for evolution, and given that extension from this position has also been noted to cause a loss of activity in EcHPPK,⁵³ it is anticipated that increased selectivity may accompany derivatization.

It is well established that broad spectrum antibiotics have detrimental effects on the host microbiome,⁵⁵ which plays an essential role in human health. A selective agent would reduce these effects, with an associated reduction in the development

of antibiotic resistance⁵⁶ and risk of secondary infections.^{55,57,58} The mechanistic insight revealed by the benzylated 8MG compounds reported in this work offers a rational course for the development of higher affinity and increasingly selective inhibitors of SaHPPK as leads for novel antimicrobials.

EXPERIMENTAL SECTION

Preparation of EcHPPK Enzyme. A pET28A plasmid containing the synthesized EcHPPK sequence (Geneart) was cloned with an N-terminal hexahistidine tag and a TEV cleavage site. *E. coli* BL21 (DE3) cells were transformed and grown overnight in 5 mL of 2xYT medium supplemented with $50 \mu\text{g}\cdot\text{mL}^{-1}$ kanamycin for selection. The overnight culture was then subcultured into fresh 2xYT with growth at 310 K for ~ 2 h until an OD_{600} of 0.5–0.7 was reached. Isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM at 298 K for 18 h. The cultures were centrifuged at 5000 rpm for 10 min and the cells resuspended in 50 mL of HEPES buffer, pH 8.0, 5% glycerol. An EDTA-free Complete protease-inhibitor cocktail tablet (Roche) was added together with lysozyme to a final concentration of $0.4 \text{ mg}\cdot\text{mL}^{-1}$. After 30 min, the cells were sonicated and the cell debris was removed by centrifugation at 18 000 rpm at 277 K for 30 min. The supernatant was filtered ($0.45 \mu\text{m}$ filter) and loaded onto a Ni-NTA IMAC column (Qiagen). Unbound protein was washed off with 10 mM imidazole in 50 mM HEPES buffer, pH 8.0, 0.15 M NaCl, 5% glycerol, 1 mM DTT. Protein was eluted from the column with a 500 mM imidazole-containing variant of the above buffer. To cleave the His-tag, 1 mg of TEV protease was added to the sample and left overnight at 4 °C for 16 h. The sample was reloaded onto the Ni-NTA IMAC column to separate cleaved from uncleaved protein. Each sample was further purified using a Superdex 75 size-

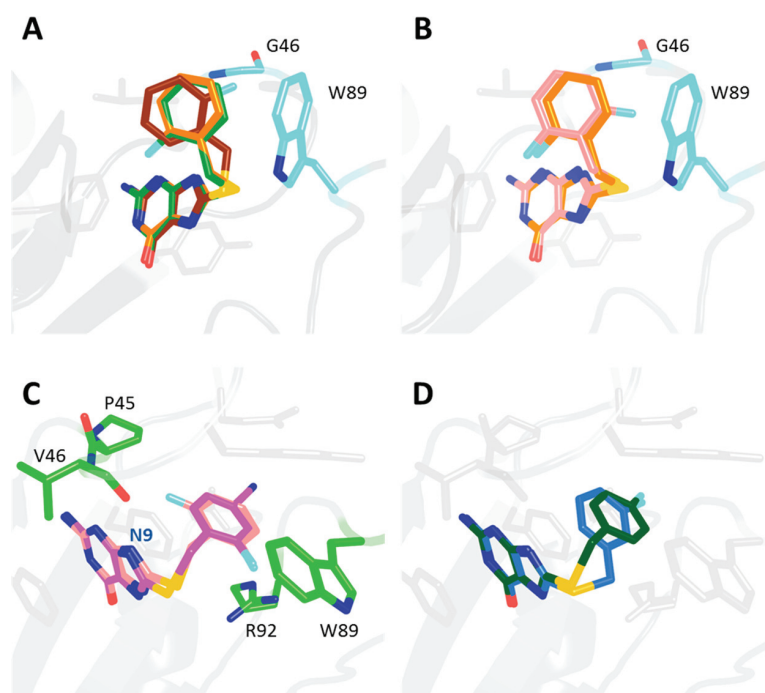


Figure 8. (A) **18** bound to ECHPPK as obtained from the crystal structure (orange), native ligand docking (green), or cross-docking (brown). (B) Docked model of **41** (pink) to ECHPPK. **18** from its ECHPPK crystal complex is displayed in orange. (C) Docked model of **41** (pink) to SaHPPK. **50** from its SaHPPK crystal complex is displayed in purple. “N9” indicates position N9 of the guanine ring. (D) Docked model of **25** (blue) to SaHPPK. **34** from its SaHPPK crystal complex is displayed in dark green.

exclusion 26/60 column (GE Healthcare) using buffer conditions reported in ref 19, i.e., 20 mM Tris-HCl buffer, pH 8.0, 100 mM NaCl. Fractions were analyzed using a 15% SDS-PAGE gel with Coomassie staining. His-tagged and cleaved protein samples were concentrated to 2.2 mg·mL⁻¹ and 6.6 mg·mL⁻¹ for use in SPR and X-ray crystallographic experiments, respectively. All samples were snap frozen and stored at 193 K.

Additional Enzyme Preparation. Samples of isotopically labeled SaHPPK for NMR spectroscopy were expressed as described previously³¹ with purification as per ref 30. SaHPPK samples for crystallography and SPR were expressed and purified as per ref 30.

NMR Spectroscopy. All NMR experiments were recorded at 295 K on a Bruker Avance 600 MHz NMR spectrometer equipped with a cryoprobe and Z axis gradient. Triple resonance NMR experiments were performed on a sample of ~0.25 mM ¹⁵N/¹³C-labeled SaHPPK dissolved in a 90%/10% H₂O/D₂O 50 mM HEPES buffer, 1% sorbitol, and 2% DMSO-*d*₆ at pH 8.0 in the presence of 10 mM MgCl₂ and saturating amounts of AMPCPP (1 mM). Titrations of compounds **17**, **40**, **41**, or **50** to saturation were performed from a 50 mM stock dissolved in DMSO-*d*₆. Backbone assignments were obtained using the HNCA experiment. The experiment used a WATERGATE sequence for solvent suppression. Spectra were processed using NMRPipe⁵⁹ and analyzed with XEASY⁶⁰ or SPARKY.⁶¹ 2D ¹⁵N FAST-HSQC experiments were typically acquired with $t_{1\max}({}^{15}\text{N}) = 51\text{--}62$ ms and $t_{2\max}({}^1\text{H}) = 142$ ms, whereas triple resonance experiments were acquired with $t_{1\max}({}^{15}\text{N}) = 23.3$ ms, $t_{2\max}({}^{13}\text{C}) = 10.4$ ms, $t_{2\max}({}^1\text{H}) = 15.1$ ms, and $t_{3\max}({}^1\text{H}) = 142$ ms.

Crystallization and X-ray Structure Determination. Crystallization experiments were performed at the CSIRO Collaborative Crystallisation Centre. The sitting-drop vapor diffusion method was used at either 281 or 293 K with droplets consisting of 150 nL of protein solution and 150 nL of reservoir solution and using a reservoir volume of 50 μL. Crystals of the **13**/AMPCPP/SaHPPK complex were obtained from a solution containing 0.186 M sodium nitrate and 18.4% (w/v) PEG 3000, with addition of 2 mM MgCl₂, 1 mM AMPCPP, and 1 mM inhibitor. Crystals of SaHPPK/AMPCPP complexed with compounds **34**, **35**, **40**, and **50** were obtained from

solutions containing approximately 0.1 M TrisCl, pH 8.5, 0.2 M MgCl₂, and 20% PEG 8000 with addition of 0.05 M sodium thiocyanate, 1 mM AMPCPP, and 1 mM inhibitor. The SaHPPK protein was present at a concentration of 6.9 mg·mL⁻¹. For the ECHPPK complexes, crystallization conditions similar to those used by Yun et al.⁵³ were implemented. The complex of **13**/AMPCPP/ECHPPK was therefore grown in 20% (w/v) PEG 4000, 0.1 M TrisCl, and 0.172 M CaCl₂. The crystals of AMPCPP/ECHPPK in complex with **18**, **24**, **34**, **35**, and **50** were obtained in approximately 0.1 M HEPES-NaOH, pH 7.5, 0.2 M CaCl₂, and 25–30% PEG 4000. All ECHPPK crystallization solutions contained 2 mM MgCl₂, 1 mM AMPCPP, and 1 mM of inhibitor. Data were collected at the MX-2 beamline of the Australian Synchrotron (see Tables S1 and S2 for data collection and refinement statistics) using an ADSC Quantum 315 detector, 360 frames obtained with a 1° oscillation angle for a complete data set. These data were indexed using XDS⁶² and scaled using SCALA.⁶³ The ECHPPK and SaHPPK PDB structures 1QON⁶⁴ and 4CWB⁵³ respectively were used to solve the initial phases of the complexes by molecular replacement using Phaser.⁶⁵ Refinement was performed using REFMAC5,⁶⁶ and the electron density maps were visualized in Coot.⁶⁷ After several rounds of manual rebuilding, ligands and water molecules were added and the models further refined. Crystallographic information files (CIFs) for the ligands were generated using AFITT.⁶⁸ PDB_REDO⁶⁹ was used to optimize refinement parameters. In the ECHPPK complexes, ~0.2 M CaCl₂ is present in the crystallization conditions; the high-quality density strongly indicated that Ca²⁺ ions had replaced Mg²⁺ ions at the active site, and the structures were refined accordingly. As noted,⁵³ this crystallization artifact does not, however, appear to perturb compound binding. Data collection and refinement statistics are summarized in Table 4 and fully presented in Supporting Information Tables S1 and S2.

Surface Plasmon Resonance (SPR). All SPR experiments were performed using a ProteOn XPR36 (Bio-Rad) biosensor with a GLH chip. Streptavidin was coupled to the chip surface at 310 K in HBS-EP + running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 50 μM EDTA, 0.05% [v/v] Tween-20) after activation with a 1:1 mixture of

NHS/EDC (*N*-hydroxysuccinimide/*N*-ethyl-*N'*-(3-diethylaminopropyl)carbodiimide). Protein was diluted to 100 $\mu\text{g}\cdot\text{mL}^{-1}$ at pH 4 and injected for 6 min. The surface was then blocked with 1 M ethanolamine, pH 8.0, with an immobilization of 15 000–16 000 RU for each flow cell. EcHPPK and SaHPPK were minimally biotinylated using NHS-activated biotin and maleimide-activated biotin, respectively. HPPK immobilization was performed at 293 K in 50 mM HEPES, 150 mM NaCl, 10 mM MgCl_2 , 3% DMSO, 0.05% Tween, 2.5 mM DTT, pH 8.0. Proteins were injected over streptavidin-coupled flow-cells for 25 min. Immobilization levels for EcHPPK and SaHPPK were ~ 4900 and ~ 3800 RU, respectively. Biotin was subsequently injected over HPPK-bound flow-cells to block remaining unbound streptavidin. All SPR binding experiments were performed at 293 K in the above immobilization buffer but with the addition of 1 mM ATP. Analytes were serially diluted (3-fold) in SPR binding buffer and injected for 60 s contact time at 30 $\mu\text{L}\cdot\text{min}^{-1}$, then allowed to dissociate for 60 s. Each analyte titration was performed in duplicate or greater. Binding sensorgrams were processed, solvent-corrected, and double-referenced using ProteOn Manager software, version 3.1.0 (Bio-Rad). SPR binding analysis of several of the compounds investigated in this study revealed dissociation rates that were not sufficiently slow to allow global fitting to a kinetic binding model, for which the k_d (dissociation rate constant) must typically be $<0.5\text{ s}^{-1}$ for SPR instruments to be able to capture sufficient data points during the dissociation phase. Therefore, to determine binding affinities (K_D values), responses at equilibrium for each analyte were fitted to a 1:1 steady-state affinity model available within the ProteOn Manager software, version 3.1.0 (Bio-Rad). Compounds were tested from a top concentration of 3–300 μM , ideally ~ 10 times the K_D . Where solubility or affinity was limiting in this regard (the top concentration being less than 3 times the K_D), compounds were globally fitted with a compound that was approaching saturation to the protein. Sensorgrams and binding isotherms are displayed in Supporting Information Figure S1.

Molecular Modeling. Molecular modeling was performed using the Schrödinger Suite 2015 (www.schrodinger.com) through the Maestro interface (Maestro, version 10.3, Schrödinger, LLC, New York, 2015).⁷⁰ Protein preparation was performed with the Protein Preparation Wizard workflow implemented by Schrödinger (Epik, version 3.3, Schrödinger, LLC, New York, 2015),⁷¹ with deletion of all waters. In order to eliminate any bond length or bond angle biases in the structures, compounds were subjected to a full minimization prior to docking using LigPrep (LigPrep, version 3.5, Schrödinger, LLC, New York, 2015).⁷² Docking was carried out with Glide,⁵⁴ version 6.8, using “standard precision” (SP) or “extra precision” (XP) mode as stated.

Compound Procurement and Analysis. 8MG and 14 were purchased from Sigma-Aldrich and TimTec, respectively. All other compounds were synthesized as described below. In all cases, ^1H NMR spectra were recorded on a 400 MHz Bruker NMR spectrometer, and chemical shifts were referenced to the residual proteo-solvent peak. NMR spectra were run in DMSO unless otherwise stated. Chemical shifts are reported as δ values in parts per million (ppm), and coupling constants (J) are reported in hertz (Hz). Exchangeable protons at the N1, N3, N7, N9, or O6 position can potentially be observed as broad signals at high frequency in the ^1H NMR spectrum, and their observation depends on the protonation state and potential intermolecular interactions of the guanine.⁷³ In some cases, 0.1 μL of glacial acetic acid was added to shift the equilibrium to a single species for ^1H NMR analyses. The broad signals at >10 ppm in the ^1H NMR spectra of the compounds below are reported if observed but have not been specifically assigned to NH or OH. Liquid chromatography–mass spectrometry (LCMS) was acquired on an Agilent UHPLC/MS (1260/6120) system using a Poroshell 120 EC-C18 column (3.0 mm \times 50 mm, 2.7 μm) with a flow rate of 0.5 $\text{mL}\cdot\text{min}^{-1}$. Elution was achieved with standard HPLC buffers (buffer A, 99.9% $\text{H}_2\text{O}/0.1\%$ formic acid; buffer B, 99.9% $\text{CH}_3\text{CN}/0.1\%$ formic acid) using a gradient from 5% B/95% A to 100% B over 5 min. Preparative reverse-phase HPLC was performed on an Agilent 1260 preparative HPLC using an Alltima C8 column

(250 mm \times 22 mm, 5 μm) with a flow rate of 20 $\text{mL}\cdot\text{min}^{-1}$. Elution was achieved using preparative HPLC buffers (buffer C, 99.9% $\text{H}_2\text{O}/0.1\%$ TFA; buffer D, 99.9% $\text{CH}_3\text{CN}/0.1\%$ TFA) using a gradient from 5% D/95% C to 95% D over 20 min. Final compounds were determined to be $>95\%$ pure by analytical HPLC and ^1H NMR.

General Procedure for the Alkylation of 8MG. 8MG (0.1 g, 0.55 mmol) was dissolved in 0.5 M NaOH solution (3 mL), and to the stirring solution alkylating reagent (0.6 mmol) was added dropwise. The reaction was allowed to proceed overnight at rt. The reaction mixture was then made mildly acidic by the dropwise addition of acetic acid resulting in the precipitation of an off-white amorphous solid. The solid was filtered and dried to give the title compound with yields ranging from 9% to 99%. Compounds 5, 7, 10, 37, 38, 43, and 44 were further purified by preparative reverse-phase HPLC (vide supra). Fractions containing pure products were lyophilized to yield the title compounds as white or off-white solids.

2-Amino-8-((2-hydroxyethyl)thio)-1,9-dihydro-6H-purin-6-one (1). Yield 36%. ^1H NMR (400 MHz, DMSO) δ 6.35 (s, 2H, NH_2), 3.63 (t, $J = 6.5$ Hz, 2H, CH_2O), 3.18 (t, $J = 6.5$ Hz, 2H, CH_2S). LCMS: 228.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((3-hydroxypropyl)thio)-1,9-dihydro-6H-purin-6-one (2). Yield 76%. ^1H NMR (400 MHz, DMSO) δ 6.62 (s, 2H, NH_2), 3.48 (t, $J = 6.1$ Hz, 2H, CH_2O), 3.13 (t, $J = 7.0$ Hz, 2H, CH_2S), 1.80–1.73 (m, 2H, CH_2). LCMS: 242.0 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2,3-dihydroxypropyl)thio)-1,9-dihydro-6H-purin-6-one (3). Yield 47%. ^1H NMR (400 MHz, DMSO) δ 6.39 (s, 2H, NH_2), 3.69 (m, 1H, CH), 3.37 (m, 2H, CH_2O), 3.28 (dd, $J = 13.5$, 4.6 Hz, 1H, CH_2), 3.11 (dd, $J = 13.5$, 6.6 Hz, 1H, CH_2). LCMS: 258.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((3,3,3-trifluoro-2-hydroxypropyl)thio)-1,9-dihydro-6H-purin-6-one (4). Yield 31%. ^1H NMR (400 MHz, DMSO) δ 10.90–10.20 (bs, 1H), 6.23 (s, 2H, NH_2), 4.26 (m, 1H, CH), 3.47 (dd, $J = 13.9$, 2.9 Hz, 1H, CH_2), 3.14 (dd, $J = 13.9$, 9.5 Hz, 1H, CH_2). LCMS: 296.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2-methoxyethyl)thio)-1,9-dihydro-6H-purin-6-one (5). Yield 60%. ^1H NMR (400 MHz, DMSO) δ 10.73 (bs, 1H), 6.47 (bs, 2H, NH_2), 3.57 (t, $J = 6.3$ Hz, 2H, CH_2), 3.32 (t, $J = 6.3$ Hz, 2H, CH_2), 3.26 (s, 3H, CH_3). LCMS: 242.1 $[\text{M} + \text{H}]^+$.

8-((2-(1,3-Dioxan-2-yl)ethyl)thio)-2-amino-1,9-dihydro-6H-purin-6-one (6). Yield 35%. ^1H NMR (400 MHz, DMSO) δ 12.51 (bs, 1H), 10.53 (bs, 1H), 6.25 (s, 2H, NH_2), 4.63 (t, $J = 5.0$ Hz, 1H, CH), 3.99 (dd, $J = 10.8$, 4.9 Hz, 2H, cyclic CH_2O), 3.69 (td, $J = 12.3$, 2.2 Hz, 2H, cyclic CH_2O), 3.10 (t, $J = 7.4$ Hz, 2H, CH_2S), 1.92–1.77 (m, 3H, CH_2 and cyclic CH_2), 1.33 (d, $J = 13.4$ Hz, 1H, cyclic CH_2). LCMS: 298.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2-(2-oxopyrrolidin-1-yl)ethyl)thio)-1,9-dihydro-6H-purin-6-one (7). Yield 53%. ^1H NMR (400 MHz, DMSO) δ 10.89 (bs, 1H), 6.61 (s, 2H, NH_2), 3.48 (t, $J = 6.5$ Hz, 2H, CH_2), 3.38 (t, $J = 7.0$ Hz, 2H, CH_2), 3.30 (t, $J = 6.5$ Hz, 2H, CH_2), 2.18 (t, $J = 8.1$ Hz, 2H, CH_2), 1.88 (m, 2H, CH_2). LCMS: 295.1 $[\text{M} + \text{H}]^+$.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N-methylacetamide (8). Yield 37%. ^1H NMR (400 MHz, DMSO) δ 8.61 (bs, 1H, NH), 6.51 (s, 2H, NH_2), 3.73 (s, 2H, CH_2), 2.59 (d, $J = 4.5$ Hz, 3H, CH_3). LCMS: 255.1 $[\text{M} + \text{H}]^+$.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N-isobutylacetamide (9). Yield 40%. ^1H NMR (400 MHz, DMSO) δ 8.54 (s, 1H, NH), 6.49 (s, 2H, NH_2), 3.77 (s, 2H, CH_2S), 2.89 (app t, $J = 6.2$ Hz, 2H, CH_2NH), 1.65 (m, 1H, CH), 0.80 (d, $J = 6.7$ Hz, 6H, CH_3). LCMS: 297.1 $[\text{M} + \text{H}]^+$.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N,N-dimethylacetamide (10). Yield 10%. ^1H NMR (400 MHz, DMSO) δ 10.95 (bs, 1H), 6.68 (bs, 2H, NH_2), 4.23 (s, 2H, CH_2S), 3.38 (q, 3H, CH_3), 2.85 (s, 3H, CH_3). LCMS: 269.1 $[\text{M} + \text{H}]^+$.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N,N-diethylacetamide (11). Yield 44%. ^1H NMR (400 MHz, DMSO) δ 10.57 (bs, 1H), 6.30 (s, 2H, NH_2), 4.23 (s, 2H, CH_2S), 3.38 (q, $J = 7.1$ Hz, 2H, CH_2N), 3.28 (q, $J = 7.1$ Hz, 2H, CH_2N), 1.16 (t, $J = 7.1$ Hz, 3H, CH_3), 1.01 (d, $J = 7.1$ Hz, 3H, CH_3). LCMS: 297.1 $[\text{M} + \text{H}]^+$.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N-ethyl-N-methylpropanamide (12). Yield 21%. ^1H NMR (400 MHz, DMSO) δ 10.89 (bs, 1H), 6.62 (bs, 2H, NH_2), 4.75 (m, 1H, CH),

3.40–3.24 (m, 2H, CH₂), 3.05 (s, 1.5H, NCH₃), 2.83 (s, 1.3H, NCH₃), 1.47 (m, 3H, CH₃CH), 1.15 (t, *J* = 7.1 Hz, 1.3H, CH₃CH₂), 0.99 (t, *J* = 7.1 Hz, 1.5H, CH₃CH₂). LCMS: 297.1 [M + H]⁺. Two sets of signals were observed due to the planar nature of the amide bond having either the *E* or *Z* isomer in slow exchange on the NMR time scale, adjacent to the *R* or *S* stereogenic carbon.

2-Amino-8-((2-(4-bromophenyl)-2-oxoethyl)thio)-1,9-dihydro-6H-purin-6-one (13). Yield 72%. ¹H NMR (400 MHz, DMSO) δ 10.68 (bs, 1H), 7.96 (d, *J* = 8.5 Hz, 2H, ArH), 7.77 (d, *J* = 8.5 Hz, 2H, ArH), 6.31 (s, 2H, NH₂), 4.84 (s, 2H, CH₂). LCMS: 379.9 [M + H]⁺.

8-((2-([1,1'-Biphenyl]-4-yl)-2-oxoethyl)thio)-2-amino-1,9-dihydro-6H-purin-6-one (15). Yield 9%. ¹H NMR (400 MHz, DMSO) δ 10.95 (bs, 1H), 8.12 (d, *J* = 8.4 Hz, 2H, ArH), 7.82 (d, *J* = 8.5 Hz, 2H, ArH), 7.76 (m, 2H, ArH), 7.50 (m, 2H, ArH), 7.43 (m, 1H, ArH), 6.39 (s, 2H, NH₂), 4.80 (s, 2H, CH₂). LCMS: 378.1 [M + H]⁺.

2-Amino-8-((2-oxo-2-phenylethyl)thio)-1,9-dihydro-6H-purin-6-one (16). Yield 50%. ¹H NMR (400 MHz, DMSO) δ 10.91 (bs, 1H), 8.02 (m, 2H, ArH), 7.67 (t, *J* = 7.4 Hz, 1H, ArH), 7.55 (t, *J* = 7.7 Hz, 2H, ArH), 6.45 (s, 2H, NH₂), 4.86 (s, 2H, CH₂). LCMS: 302.1 [M + H]⁺.

2-Amino-8-(benzylthio)-1,9-dihydro-6H-purin-6-one (17). Yield 36%. ¹H NMR (400 MHz, DMSO) δ 11.03 (bs, 1H), 7.37 (m, 2H, ArH), 7.29 (m, 2H, ArH), 7.24 (m, 1H, ArH), 6.45 (s, 2H, NH₂), 4.38 (s, 2H, CH₂). LCMS: 274.1 [M + H]⁺.

2-Amino-8-((2-fluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (18). Yield 66%. ¹H NMR (400 MHz, DMSO) δ 12.54 (bs, 1H), 10.55 (bs, 1H), 7.40 (t, *J* = 7.4 Hz, 1H, ArH), 7.30 (m, 1H, ArH), 7.17 (m, 1H, ArH), 7.11 (m, 1H, ArH), 6.28 (s, 2H, NH₂), 4.39 (s, 2H, CH₂). LCMS: 292.1 [M + H]⁺.

2-Amino-8-((2-chlorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (19). Yield 71%. ¹H NMR (400 MHz, DMSO) δ 12.55 (bs, 1H), 10.54 (bs, 1H), 7.46 (m, 2H, ArH), 7.29 (m, 2H, ArH), 6.34 (s, 2H, NH₂), 4.45 (s, 2H, CH₂). LCMS: 308.0 [M + H]⁺.

2-Amino-8-((2-bromobenzyl)thio)-1,9-dihydro-6H-purin-6-one (20). Yield 79%. ¹H NMR (400 MHz, DMSO) δ 12.59 (bs, 1H), 10.65 (bs, 1H), 7.63 (d, *J* = 7.8 Hz, 1H, ArH), 7.48 (d, *J* = 7.3 Hz, 1H, ArH), 7.31 (t, *J* = 7.3 Hz, 1H, ArH), 7.21 (t, *J* = 7.2 Hz, 1H, ArH), 6.32 (s, 2H, NH₂), 4.46 (s, 2H, CH₂). LCMS: 351.9 [M + H]⁺.

2-Amino-8-((2-methylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (21). Yield 50%. ¹H NMR (400 MHz, DMSO) δ 12.51 (bs, 1H), 10.52 (bs, 1H), 7.28 (d, *J* = 6.4 Hz, 1H, ArH), 7.17 (m, 2H, ArH), 7.10 (m, 1H, ArH), 6.31 (s, 2H, NH₂), 4.38 (s, 2H, CH₂), 2.35 (s, 3H, CH₃). LCMS: 288.1 [M + H]⁺.

2-Amino-8-((2-(trifluoromethyl)benzyl)thio)-1,9-dihydro-6H-purin-6-one (22). Yield 52%. ¹H NMR (400 MHz, DMSO) δ 12.58 (bs, 1H), 10.55 (bs, 1H), 7.73 (d, *J* = 7.8 Hz, 1H, ArH), 7.62 (app t, *J* = 6.3 Hz, 2H, ArH), 7.50 (m, 1H, ArH), 6.31 (s, 2H, NH₂), 4.54 (s, 2H, CH₂). LCMS: 342.1 [M + H]⁺.

2-Amino-8-((2-nitrobenzyl)thio)-1,9-dihydro-6H-purin-6-one (23). Yield 47%. ¹H NMR (400 MHz, DMSO) δ 12.55 (bs, 1H), 10.54 (bs, 1H), 8.05 (d, *J* = 7.8 Hz, 1H, ArH), 7.66 (m, 2H, ArH), 7.54 (m, 1H, ArH), 6.27 (s, 2H, NH₂), 4.66 (s, 2H, CH₂). LCMS: 319.0 [M + H]⁺.

2-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)benzonitrile (24). Yield 57%. ¹H NMR (400 MHz, DMSO) δ 12.57 (bs, 1H), 10.55 (bs, 1H), 7.83 (d, *J* = 7.6 Hz, 1H, ArH), 7.62 (m, 2H, ArH), 7.46 (td, *J* = 7.7, 1.2 Hz, 1H, ArH), 6.31 (s, 2H, NH₂), 4.52 (s, 2H, CH₂). LCMS: 299.1 [M + H]⁺.

2-Amino-8-((3-fluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (25). Yield 64%. ¹H NMR (400 MHz, DMSO) δ 10.78 (bs, 1H), 7.33 (m, 1H, ArH), 7.21 (m, 2H, ArH), 7.06 (m, 1H, ArH), 6.37 (s, 2H, NH₂), 4.39 (s, 2H, CH₂). LCMS: 292.1 [M + H]⁺.

2-Amino-8-((3-chlorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (26). Yield 67%. ¹H NMR (400 MHz, DMSO) δ 12.49 (bs, 1H), 10.52 (bs, 1H), 7.44 (s, 1H, ArH), 7.31 (m, 3H, ArH), 6.29 (bs, 2H, NH₂), 4.38 (s, 2H, CH₂). LCMS: 308.0 [M + H]⁺.

2-Amino-8-((3-bromobenzyl)thio)-1,9-dihydro-6H-purin-6-one (27). Yield 71%. ¹H NMR (400 MHz, DMSO) δ 10.89 (bs, 1H), 7.58 (s, 1H, ArH), 7.43 (d, *J* = 7.9 Hz, 1H, ArH), 7.36 (d, *J* = 7.7 Hz,

1H, ArH), 7.25 (app t, *J* = 7.8 Hz, 1H, ArH), 6.43 (s, 2H, NH₂), 4.37 (s, 2H, CH₂). LCMS: 351.9 [M + H]⁺.

2-Amino-8-((3-methylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (28). Yield 99%. ¹H NMR (400 MHz, DMSO) δ 12.49 (bs, 1H), 10.53 (bs, 1H), 7.16 (m, 3H, ArH), 7.05 (d, *J* = 7.2 Hz, 1H, ArH), 6.28 (s, 2H, NH₂), 4.34 (s, 2H, CH₂), 2.26 (s, 3H, CH₃). LCMS: 288.1 [M + H]⁺.

2-Amino-8-((3-methoxybenzyl)thio)-1,9-dihydro-6H-purin-6-one (29). Yield 72%. ¹H NMR (400 MHz, DMSO) δ 12.50 (bs, 1H), 10.52 (bs, 1H), 7.20 (m, 1H, ArH), 6.92 (m, 2H, ArH), 6.81 (m, 1H, ArH), 6.30 (bs, 2H, NH₂), 4.35 (s, 2H, CH₂), 3.70 (s, 3H, CH₃). LCMS: 304.0 [M + H]⁺.

2-Amino-8-((3-(trifluoromethoxy)benzyl)thio)-1,9-dihydro-6H-purin-6-one (30). Yield 32%. ¹H NMR (400 MHz, DMSO) δ 12.53 (bs, 1H), 10.53 (bs, 1H), 7.40 (m, 3H, ArH), 7.23 (d, *J* = 7.7 Hz, 1H, ArH), 6.28 (s, 2H, NH₂), 4.43 (s, 2H, CH₂). LCMS: 358.1 [M + H]⁺.

2-Amino-8-((3-nitrobenzyl)thio)-1,9-dihydro-6H-purin-6-one (31). Yield 69%. ¹H NMR (400 MHz, DMSO) δ 10.59 (bs, 1H), 8.26 (s, 1H, ArH), 8.09 (m, 1H, ArH), 7.82 (d, *J* = 7.8 Hz, 1H, ArH), 7.59 (t, *J* = 7.9 Hz, 1H, ArH), 6.27 (s, 2H, NH₂), 4.51 (s, 2H, CH₂). LCMS: 319.0 [M + H]⁺.

3-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)benzonitrile (32). Yield 71.4%. ¹H NMR (400 MHz, DMSO) δ 12.51 (bs, 1H), 10.52 (bs, 1H), 7.82 (s, 1H, ArH), 7.71 (m, 2H, ArH), 7.51 (t, *J* = 7.8 Hz, 1H, ArH), 6.31 (bs, 2H, NH₂), 4.42 (s, 2H, CH₂). LCMS: 299.1 [M + H]⁺.

3-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)benzoic Acid (33). Yield 91%. ¹H NMR (400 MHz, DMSO) δ 11.81 (bs, 1H), 7.98 (s, 1H, ArH), 7.77 (dt, *J* = 7.6, 1.2 Hz, 1H, ArH), 7.37 (d, *J* = 8.0 Hz, 1H, ArH), 7.25 (t, *J* = 7.6 Hz, 1H, ArH), 6.84 (s, 2H, NH₂), 4.42 (s, 2H, CH₂). LCMS: 318.1 [M + H]⁺.

2-Amino-8-((4-fluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (34). Yield 56%. ¹H NMR (400 MHz, DMSO) δ 12.49 (bs, 1H), 10.52 (bs, 1H), 7.40 (m, 2H, ArH), 7.12 (m, 2H, ArH), 6.30 (bs, 2H, NH₂), 4.37 (s, 2H, CH₂). LCMS: 292.1 [M + H]⁺.

2-Amino-8-((4-chlorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (35). Yield 42%. ¹H NMR (400 MHz, DMSO) δ 12.58 (bs, 1H), 10.60 (bs, 1H), 7.36 (m, 4H, ArH), 6.29 (s, 2H, NH₂), 4.37 (s, 2H, CH₂). LCMS: 308.0 [M + H]⁺.

2-Amino-8-((4-bromobenzyl)thio)-1,9-dihydro-6H-purin-6-one (36). Yield 63%. ¹H NMR (400 MHz, DMSO) δ 12.50 (bs, 1H), 7.49 (d, *J* = 8.4 Hz, 2H, ArH), 7.31 (d, *J* = 7.8 Hz, 2H, ArH), 6.30 (bs, 2H, NH₂), 4.35 (s, 2H, CH₂). LCMS: 351.9 [M + H]⁺.

2-Amino-8-((4-methoxybenzyl)thio)-1,9-dihydro-6H-purin-6-one (37). Yield 15%. ¹H NMR (400 MHz, DMSO) δ 10.79 (bs, 1H), 7.28 (d, *J* = 8.7 Hz, 2H, ArH), 6.86 (d, *J* = 8.7 Hz, 2H, ArH), 6.55 (bs, 2H, NH₂), 4.36 (s, 2H, CH₂), 3.72 (s, 3H, CH₃). LCMS: 304.0 [M + H]⁺.

2-Amino-8-((4-(trifluoromethyl)benzyl)thio)-1,9-dihydro-6H-purin-6-one (38). Yield 48%. ¹H NMR (400 MHz, DMSO) δ 10.78 (bs, 1H), 7.67 (d, *J* = 8.1 Hz, 2H, ArH), 7.59 (d, *J* = 8.1 Hz, 2H, ArH), 6.51 (bs, 2H, NH₂), 4.49 (s, 2H, CH₂). LCMS: 342.1 [M + H]⁺.

2-Amino-8-((4-(trifluoromethoxy)benzyl)thio)-1,9-dihydro-6H-purin-6-one (39). Yield 63%. ¹H NMR (400 MHz, DMSO) δ 12.56 (bs, 1H), 10.59 (bs, 1H), 7.49 (d, *J* = 8.7 Hz, 2H, ArH), 7.30 (m, 2H, ArH), 6.28 (bs, 2H, NH₂), 4.42 (s, 2H, CH₂). LCMS: 358.1 [M + H]⁺.

4-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)benzonitrile (40). Yield 73%. ¹H NMR (400 MHz, DMSO) δ 10.63 (bs, 1H), 7.77 (d, *J* = 8.4 Hz, 2H, ArH), 7.55 (d, *J* = 8.4 Hz, 2H, ArH), 6.30 (s, 2H, NH₂), 4.45 (s, 2H, CH₂). LCMS: 300.1 [M + H]⁺.

2-Amino-8-((2,6-difluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (41). Yield 68%. ¹H NMR (400 MHz, DMSO) δ 12.59 (bs, 1H), 10.54 (bs, 1H), 7.38 (m, 1H, ArH), 7.08 (m, 2H, ArH), 6.32 (s, 2H, NH₂), 4.35 (s, 2H, CH₂). LCMS: 310.1 [M + H]⁺.

2-Amino-8-((3,5-dimethylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (42). Yield 81%. ¹H NMR (400 MHz, DMSO) δ 10.66 (bs, 1H), 6.95 (s, 2H, ArH), 6.87 (s, 1H, ArH), 6.34 (s, 2H, NH₂), 4.31 (s, 2H, CH₂), 2.21 (s, 6H, CH₃). LCMS: 302.1 [M + H]⁺.

2-Amino-8-((3,5-bis(trifluoromethyl)benzyl)thio)-1,9-dihydro-6H-purin-6-one (43). Yield 74%. ^1H NMR (400 MHz, DMSO) δ 10.69 (bs, 1H), 8.06 (s, 2H, ArH), 7.97 (s, 1H, ArH), 6.45 (bs, 2H, NH_2), 4.55 (s, 2H, CH_2). LCMS: 410.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2,3-dimethylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (44). Yield 43%. ^1H NMR (400 MHz, DMSO) δ 12.50 (bs, 1H), 10.54 (bs, 1H), 7.12 (d, $J = 7.2$ Hz, 1H, ArH), 7.07 (d, $J = 7.3$ Hz, 1H, ArH), 6.99 (app t, $J = 7.5$ Hz, 1H, ArH), 6.31 (s, 2H, NH_2), 4.40 (s, 2H, CH_2), 2.24 (s, 6H, CH_3). LCMS: 302.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2,5-dimethylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (45). Yield 76%. ^1H NMR (400 MHz, DMSO) δ 12.49 (bs, 1H), 10.50 (bs, 1H), 7.10 (s, 1H, ArH), 7.06 (d, $J = 7.7$ Hz, 1H, ArH), 6.98 (d, $J = 7.7$ Hz, 1H, ArH), 6.31 (s, 2H, NH_2), 4.34 (s, 2H, CH_2), 2.30 (s, 3H, CH_3), 2.20 (s, 3H, CH_3). LCMS: 302.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2-fluoro-3-methylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (46). Yield 28%. ^1H NMR (400 MHz, DMSO) δ 12.54 (bs, 1H), 10.54 (bs, 1H), 7.18 (m, 2H, ArH), 7.00 (t, $J = 7.6$ Hz, 1H, ArH), 6.31 (s, 2H, NH_2), 4.38 (s, 2H, CH_2), 2.22 (d, $J = 1.9$ Hz, 3H, CH_3). LCMS: 306.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2-fluoro-5-methylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (47). Yield 55%. ^1H NMR (400 MHz, DMSO) δ 12.56 (bs, 1H), 10.58 (bs, 1H), 7.20 (d, $J = 6.2$ Hz, 1H, ArH), 7.08 (m, 2H, ArH), 6.33 (s, 2H, NH_2), 4.35 (s, 2H, CH_2), 2.21 (s, 3H, CH_3). LCMS: 306.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2,4-difluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (48). Yield 44%. ^1H NMR (400 MHz, DMSO) δ 12.54 (bs, 1H), 10.54 (bs, 1H), 7.45 (m, 1H, ArH), 7.22 (m, 1H, ArH), 7.02 (m, 1H, ArH), 6.29 (bs, 2H, NH_2), 4.36 (s, 2H, CH_2). LCMS: 310.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((2-fluoro-4-methoxybenzyl)thio)-1,9-dihydro-6H-purin-6-one (49). Yield 35%. ^1H NMR (400 MHz, DMSO) δ 12.47 (bs, 1H), 10.62 (s, 1H), 7.31 (t, $J = 8.7$ Hz, 1H, ArH), 6.81 (dd, $J = 12.1, 2.3$ Hz, 1H, ArH), 6.71 (dd, $J = 8.5, 2.3$ Hz, 1H, ArH), 6.32 (s, 2H, NH_2), 4.33 (s, 2H, CH_2), 3.73 (s, 3H, CH_3). LCMS: 322.1 $[\text{M} + \text{H}]^+$.

4-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)-3-fluorobenzonitrile (50). Yield 11%. ^1H NMR (400 MHz, DMSO) δ 12.57 (bs, 1H), 10.53 (bs, 1H), 7.83 (d, $J = 9.5$ Hz, 1H, ArH), 7.62 (m, 2H, ArH), 6.29 (bs, 2H, NH_2), 4.44 (s, 2H, CH_2). LCMS: 317.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((4-fluoro-2-methylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (51). Yield 66%. ^1H NMR (400 MHz, DMSO) δ 10.85 (bs, 1H), 7.31 (dd, $J = 8.4, 6.1$ Hz, 1H, ArH), 7.05 (dd, $J = 10.0, 2.7$ Hz, 1H, ArH), 6.93 (td, $J = 8.5, 2.7$ Hz, 1H, ArH), 6.42 (s, 2H, NH_2), 4.36 (s, 2H, CH_2), 2.36 (s, 3H, CH_3). LCMS: 306.1 $[\text{M} + \text{H}]^+$.

2-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)-5-fluorobenzonitrile (52). Yield 76%. ^1H NMR (400 MHz, DMSO) δ 10.88 (bs, 1H), 7.83 (dd, $J = 8.6, 2.5$ Hz, 1H, ArH), 7.60 (dd, $J = 8.5, 5.5$ Hz, 1H, ArH), 7.52 (td, $J = 8.6, 2.6$ Hz, 1H, ArH), 6.42 (s, 2H, NH_2), 4.49 (s, 2H, CH_2). LCMS: 317.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((3,4-difluorobenzyl)thio)-1,9-dihydro-6H-purin-6-one (53). Yield 73%. ^1H NMR (400 MHz, DMSO) δ 12.52 (bs, 1H), 10.55 (bs, 1H), 7.44 (m, 1H, ArH), 7.33 (m, 1H, ArH), 7.20 (m, 1H, ArH), 6.28 (bs, 2H, NH_2), 4.37 (s, 2H, CH_2). LCMS: 310.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((4-fluoro-3-methylbenzyl)thio)-1,9-dihydro-6H-purin-6-one (54). Yield 54%. ^1H NMR (400 MHz, DMSO) δ 10.80 (bs, 1H), 7.27 (dd, $J = 7.5, 1.8$ Hz, 1H, ArH), 7.19 (m, 1H, ArH), 7.05 (m, 1H, ArH), 6.38 (s, 2H, NH_2), 4.33 (s, 2H, CH_2), 2.18 (d, $J = 1.7$ Hz, 3H, CH_3). LCMS: 306.1 $[\text{M} + \text{H}]^+$.

5-(((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)methyl)-2-fluorobenzonitrile (55). Yield 16%. ^1H NMR (400 MHz, DMSO) δ 12.53 (bs, 1H), 10.51 (bs, 1H), 7.91 (m, 1H, ArH), 7.77 (m, 1H, ArH), 7.46 (t, $J = 9.1$ Hz, 1H, ArH), 6.23 (s, 2H, NH_2), 4.39 (s, 2H, CH_2). LCMS: 317.1 $[\text{M} + \text{H}]^+$.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.6b00002.

Sensorgrams and binding isotherms corresponding to the SPR performed (Figure S1); ligand difference ($F_o - F_c$) densities for the crystal structures (Figure S2); chemical structures of reported compounds (Figure S3); crystallographic data collection and refinement statistics (Tables S1 and S2); ligand rmsd values for native docking and cross-docking (Tables S3 and S4) (PDF)

Molecular formula strings (CSV)

Accession Codes

Ternary complexes of compounds bound to AMPCPP/EcHPPK are as follows: 13, SETP; 18, SETK; 24, SETL; 34, SETM; 35, SETN; 50, SETO. Ternary complexes of compounds bound to AMPCPP/SaHPPK are as follows: 13, SETV; 34, SETR; 35, SETS; 40, SETQ; 50, SETT.

■ AUTHOR INFORMATION

Corresponding Author

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

All crystals were grown at the C3 Crystallisation Centre at CSIRO, Parkville, Australia, and X-ray data were obtained at the MX2 beamline at the Australian Synchrotron, Victoria, Australia. We thank OpenEye for a license to use their software. We thank Olan Dolezal for his assistance with the SPR performed, Janet Newman for help with crystallization, and Bim Graham for critically proofreading the manuscript. J.B.B. was supported by an NHMRC Research Fellowship 1020411.

■ ABBREVIATIONS USED

HMDP, 6-hydroxymethyl-7,8-dihydropterin; SaHPPK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *S. aureus*; EcHPPK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *E. coli*; MRSA, methicillin-resistant *S. aureus*; 8MG, 8-mercaptopguanine; SPR, surface plasmon resonance; TEV, tobacco etch virus; SI, selectivity index; AMPCPP, α,β -methyleneadenosine 5'-triphosphate; RMSD, root-mean-square deviation; SP, standard precision; XP, extra precision

■ REFERENCES

- (1) Frieden, T. Antibiotic resistance threats in the United States, 2013. Executive summary. Centers for Disease Control and Prevention: Atlanta, GA, 2013; <http://www.cdc.gov/drugresistance/threat-report-2013/> (accessed November 12, 2015).
- (2) World Health Organization. Antimicrobial resistance: global report on surveillance 2014. <http://www.who.int/drugresistance/documents/surveillance-report/en> (accessed September 16, 2015).
- (3) Fauci, A. S.; Morens, D. M. The perpetual challenge of infectious diseases. *N. Engl. J. Med.* **2012**, *366*, 454–461.
- (4) Bax, R.; Green, S. Antibiotics: the changing regulatory and pharmaceutical industry paradigm. *J. Antimicrob. Chemother.* **2015**, *70*, 1281–1284.

- (5) Projan, S. J. Why is big Pharma getting out of antibacterial drug discovery? *Curr. Opin. Microbiol.* **2003**, *6*, 427–430.
- (6) White, A. R. Effective antibacterials: at what cost? The economics of antibacterial resistance and its control. *J. Antimicrob. Chemother.* **2011**, *66*, 1948–1953.
- (7) Rex, J. H.; Eisenstein, B. I.; Alder, J.; Goldberger, M.; Meyer, R.; Dane, A.; Friedland, I.; Knirsch, C.; Sanhai, W. R.; Tomayko, J.; Lancaster, C.; Jackson, J. A comprehensive regulatory framework to address the unmet need for new antibacterial treatments. *Lancet Infect. Dis.* **2013**, *13*, 269–275.
- (8) Guidos, R. J.; Infectious Diseases Society of America (IDSA).. Combating antimicrobial resistance: policy recommendations to save lives. *Clin. Infect. Dis.* **2011**, *52*, S397–S428.
- (9) Bermingham, A.; Derrick, J. P. The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. *BioEssays* **2002**, *24*, 637–648.
- (10) Swarbrick, J.; Iliades, P.; Simpson, J. S.; Macreadie, I. Folate biosynthesis - reappraisal of old and novel targets in the search for new antimicrobials. *Open Enzyme Inhib. J.* **2008**, *1*, 12–33.
- (11) Tari, L. W. The utility of structural biology in drug discovery. *Methods Mol. Biol.* **2012**, *841*, 1–27.
- (12) Yun, M. K.; Wu, Y.; Li, Z.; Zhao, Y.; Waddell, M. B.; Ferreira, A. M.; Lee, R. E.; Bashford, D.; White, S. W. Catalysis and sulfa drug resistance in dihydropteroate synthase. *Science* **2012**, *335*, 1110–1114.
- (13) Adra, M.; Lawrence, K. R. Trimethoprim/sulfamethoxazole for treatment of severe *Staphylococcus aureus* infections. *Ann. Pharmacother.* **2004**, *38*, 338–341.
- (14) Forgacs, P.; Wengenack, N. L.; Hall, L.; Zimmerman, S. K.; Silverman, M. L.; Roberts, G. D. Tuberculosis and trimethoprim-sulfamethoxazole. *Antimicrob. Agents Chemother.* **2009**, *53*, 4789–4793.
- (15) Martin, S. I.; Fishman, J. A. *Pneumocystis pneumonia* in solid organ transplant recipients. *Am. J. Transplant.* **2009**, *9*, S227–S233.
- (16) Pashley, T. V.; Volpe, F.; Pudney, M.; Hyde, J. E.; Sims, P. F.; Delves, C. J. Isolation and molecular characterization of the bifunctional hydroxymethyldihydropterin pyrophosphokinase-dihydropteroate synthase gene from *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* **1997**, *86*, 37–47.
- (17) Thera, M. A.; Sehdev, P. S.; Coulbaly, D.; Traore, K.; Garba, M. N.; Cissoko, Y.; Kone, A.; Guindo, A.; Dicko, A.; Beavogui, A. H.; Djimde, A. A.; Lyke, K. E.; Diallo, D. A.; Doumbo, O. K.; Plowe, C. V. Impact of trimethoprim-sulfamethoxazole prophylaxis on falciparum malaria infection and disease. *J. Infect. Dis.* **2005**, *192*, 1823–1829.
- (18) Klontz, K. C.; Singh, N. Treatment of drug-resistant *Shigella* infections. *Expert Rev. Anti-Infect. Ther.* **2015**, *13*, 69–80.
- (19) Hevener, K. E.; Yun, M. K.; Qi, J.; Kerr, I. D.; Babaoglu, K.; Hurdle, J. G.; Balakrishna, K.; White, S. W.; Lee, R. E. Structural studies of pterin-based inhibitors of dihydropteroate synthase. *J. Med. Chem.* **2010**, *53*, 166–177.
- (20) Dale, G. E.; Broger, C.; D'Arcy, A.; Hartman, P. G.; DeHoogt, R.; Jolidon, S.; Kompis, I.; Labhardt, A. M.; Langen, H.; Locher, H.; Page, M. G.; Stüber, D.; Then, R. L.; Wipf, B.; Oefner, C. A single amino acid substitution in *Staphylococcus aureus* dihydrofolate reductase determines trimethoprim resistance. *J. Mol. Biol.* **1997**, *266*, 23–30.
- (21) Frey, K. M.; Liu, J.; Lombardo, M. N.; Bolstad, D. B.; Wright, D. L.; Anderson, A. C. Crystal structures of wild-type and mutant methicillin-resistant *Staphylococcus aureus* dihydrofolate reductase reveal an alternate conformation of NADPH that may be linked to trimethoprim resistance. *J. Mol. Biol.* **2009**, *387*, 1298–1308.
- (22) Bourne, C. R.; Bunce, R. A.; Bourne, P. C.; Berlin, K. D.; Barrow, E. W.; Barrow, W. W. Crystal structure of *Bacillus anthracis* dihydrofolate reductase with the dihydrophthalazine-based trimethoprim derivative RAB1 provides a structural explanation of potency and selectivity. *Antimicrob. Agents Chemother.* **2009**, *53*, 3065–3073.
- (23) Zhao, Y.; Hammoudeh, D.; Yun, M. K.; Qi, J.; White, S. W.; Lee, R. E. Structure-based design of novel pyrimido[4,5-*c*]pyridazine derivatives as dihydropteroate synthase inhibitors with increased affinity. *ChemMedChem* **2012**, *7*, 861–870.
- (24) Hammoudeh, D. I.; Daté, M.; Yun, M. K.; Zhang, W.; Boyd, V. A.; Viacava Follis, A.; Griffith, E.; Lee, R. E.; Bashford, D.; White, S. W. Identification and characterization of an allosteric inhibitory site on dihydropteroate synthase. *ACS Chem. Biol.* **2014**, *9*, 1294–1302.
- (25) Haruki, H.; Pedersen, M. G.; Gorska, K. I.; Pojer, F.; Johnsson, K. Tetrahydrobiopterin biosynthesis as an off-target of sulfa drugs. *Science* **2013**, *340*, 987–991.
- (26) Gordon, R. J.; Lowy, F. D. Pathogenesis of Methicillin-resistant *Staphylococcus aureus* infection. *Clin. Infect. Dis.* **2008**, *46*, S350–S359.
- (27) Spaulding, A. R.; Salgado-Pabón, W.; Merriman, J. A.; Stach, C. S.; Ji, Y.; Gillman, A. N.; Peterson, M. L.; Schlievert, P. M. Vaccination against *Staphylococcus aureus* pneumonia. *J. Infect. Dis.* **2014**, *209*, 1955–1962.
- (28) Dailey, H. A.; Gerdes, S.; Dailey, T. A.; Burch, J. S.; Phillips, J. D. Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 2210–2215.
- (29) Ohlsen, K.; Lorenz, U. Novel targets for antibiotics in *Staphylococcus aureus*. *Future Microbiol.* **2007**, *2*, 655–666.
- (30) Chhabra, S.; Newman, J.; Peat, T. S.; Fernley, R. T.; Caine, J.; Simpson, J. S.; Swarbrick, J. D. Crystallization and preliminary X-ray analysis of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2010**, *66*, S75–S78.
- (31) Chhabra, S.; Dolezal, O.; Collins, B. M.; Newman, J.; Simpson, J. S.; Macreadie, I. G.; Fernley, R.; Peat, T. S.; Swarbrick, J. D. Structure of *S. aureus* HPPK and the discovery of a new substrate site inhibitor. *PLoS One* **2012**, *7*, e29444.
- (32) Chhabra, S.; Barlow, N.; Dolezal, O.; Hattarki, M. K.; Newman, J.; Peat, T. S.; Graham, B.; Swarbrick, J. D. Exploring the chemical space around 8-mercaptopurine as a route to new inhibitors of the folate biosynthesis enzyme HPPK. *PLoS One* **2013**, *8*, e59535.
- (33) Dennis, M. L.; Chhabra, S.; Wang, Z.-C.; Debono, A.; Dolezal, O.; Newman, J.; Pitcher, N. P.; Rahmani, R.; Cleary, B.; Barlow, N.; Hattarki, M.; Graham, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structure-based design and development of functionalized mercaptopurine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*. *J. Med. Chem.* **2014**, *57*, 9612–9626.
- (34) Chhabra, S.; Dolezal, B.; Hattarki, M.; Peat, T. S.; Simpson, J. S.; Swarbrick, J. D. Fragment screening on *Staphylococcus aureus* HPPK—a folate pathway target. *Aust. J. Chem.* **2013**, *66*, 1537–1543.
- (35) Blaszczyk, J.; Shi, G.; Li, Y.; Yan, H.; Ji, X. Reaction trajectory of pyrophosphoryl transfer catalyzed by 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Structure* **2004**, *12*, 467–475.
- (36) Derrick, J. P. The structure and mechanism of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Vitam. Horm.* **2008**, *79*, 411–433.
- (37) Xiao, B.; Shi, G.; Gao, J.; Blaszczyk, J.; Liu, Q.; Ji, X.; Yan, H. Unusual conformational changes in 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase as revealed by X-ray crystallography and NMR. *J. Biol. Chem.* **2001**, *276*, 40274–40281.
- (38) Lescop, E.; Lu, Z.; Liu, Q.; Xu, H.; Li, G.; Xia, B.; Yan, H.; Jin, C. Dynamics of the conformational transitions in the assembling of the Michaelis complex of a bisubstrate enzyme: a 15N relaxation study of *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Biochemistry* **2009**, *48*, 302–312.
- (39) Bermingham, A.; Bottomley, J. R.; Primrose, W. U.; Derrick, J. P. Equilibrium and kinetic studies of substrate binding to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Escherichia coli*. *J. Biol. Chem.* **2000**, *275*, 17962–17967.
- (40) Blaszczyk, J.; Li, Y.; Wu, Y.; Shi, G.; Ji, X.; Yan, H. Essential roles of a dynamic loop in the catalysis of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Biochemistry* **2004**, *43*, 1469–1477.
- (41) Shi, G.; Gong, Y.; Savchenko, A.; Zeikus, J. G.; Xiao, B.; Ji, X.; Yan, H. Dissecting the nucleotide binding properties of *Escherichia coli* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase with fluo-

rescent 3'(2)'-o-anthraniloyladenine 5'-triphosphate. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2000**, 1478, 289–299.

(42) Xiao, B.; Shi, G.; Chen, X.; Yan, H.; Ji, X. Crystal structure of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, a potential target for the development of novel antimicrobial agents. *Structure* **1999**, 7, 489–496.

(43) Hennig, M.; Dale, G. E.; D'arcy, A.; Danel, F.; Fischer, S.; Gray, C. P.; Jolidon, S.; Müller, F.; Page, M. G.; Pattison, P.; Oefner, C. The structure and function of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Haemophilus influenzae*. *J. Mol. Biol.* **1999**, 287, 211–219.

(44) Garçon, A.; Levy, C.; Derrick, J. P. Crystal structure of the bifunctional dihydroneopterin aldolase/6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Streptococcus pneumoniae*. *J. Mol. Biol.* **2006**, 360, 644–653.

(45) Lawrence, M. C.; Iliades, P.; Fernley, R. T.; Berglez, J.; Pilling, P. A.; Macreadie, I. G. The three-dimensional structure of the bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/dihydropteroyl synthase of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **2005**, 348, 655–670.

(46) Shaw, G. X.; Li, Y.; Shi, G.; Wu, Y.; Cherry, S.; Needle, D.; Zhang, D.; Tropea, J. E.; Waugh, D. S.; Yan, H.; Ji, X. Structural enzymology and inhibition of the bi-functional folate pathway enzyme HPPK-DHPS from the biowarfare agent *Francisella tularensis*. *FEBS J.* **2014**, 281, 4123–4137.

(47) Stammers, D. K.; Achari, A.; Somers, D. O.; Bryant, P. K.; Rosemond, J.; Scott, D. L.; Champness, J. N. 2.0 Å X-ray structure of the ternary complex of 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase from *Escherichia coli* with ATP and a substrate analogue. *FEBS Lett.* **1999**, 456, 49–53.

(48) Wood, H. Specific inhibition of dihydrofolate biosynthesis—a new approach to chemotherapy. In *Chemistry and Biology of Pteridines*; Walter de Gruyter: Berlin, 1975; pp 27–49.

(49) Shi, G.; Blaszczyk, J.; Ji, X.; Yan, H. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: synthesis and biochemical and crystallographic studies. *J. Med. Chem.* **2001**, 44, 1364–1371.

(50) Shi, G.; Shaw, G.; Liang, Y. H.; Subburaman, P.; Li, Y.; Wu, Y.; Yan, H.; Ji, X. Bisubstrate analogue inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: New design with improved properties. *Bioorg. Med. Chem.* **2012**, 20, 47–57.

(51) Shi, G.; Shaw, G.; Li, Y.; Wu, Y.; Yan, H.; Ji, X. Bisubstrate analog inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase: new lead exhibits a distinct binding mode. *Bioorg. Med. Chem.* **2012**, 20, 4303–4309.

(52) Shi, G.; Shaw, G.; Ji, X. Linked purine pterin HPPK inhibitors useful as antibacterial agents. U.S. Patent 9,029,344, 2015.

(53) Yun, M. K.; Hoagland, D.; Kumar, G.; Waddell, M. B.; Rock, C. O.; Lee, R. E.; White, S. W. The identification, analysis and structure-based development of novel inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Bioorg. Med. Chem.* **2014**, 22, 2157–2165.

(54) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, 49, 6177–96.

(55) Fowler, S.; Webber, A.; Cooper, B. S.; Phimister, A.; Price, K.; Carter, Y.; Kibbler, C. C.; Simpson, A. J.; Stone, S. P. Successful use of feedback to improve antibiotic prescribing and reduce *Clostridium difficile* infection: a controlled interrupted time series. *J. Antimicrob. Chemother.* **2007**, 59, 990–995.

(56) Yates, C. M.; Shaw, D. J.; Roe, A. J.; Woolhouse, M. E.; Amyes, S. G. Enhancement of bacterial competitive fitness by apramycin resistance plasmids from non-pathogenic *Escherichia coli*. *Biol. Lett.* **2006**, 2, 463–465.

(57) Lemon, K. P.; Armitage, G. C.; Relman, D. A.; Fischbach, M. A. Microbiota-targeted therapies: an ecological perspective. *Sci. Transl. Med.* **2012**, 4, 137rv5.

(58) Buffie, C. G.; Jarchum, I.; Equinda, M.; Lipuma, L.; Gbourneau, A.; Viale, A.; Ubeda, C.; Xavier, J.; Pamer, E. G. Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to *Clostridium difficile*-induced colitis. *Infect. Immun.* **2012**, 80, 62–73.

(59) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, 6, 277–293.

(60) Bartels, C.; Xia, T. H.; Billeter, M.; Güntert, P.; Wüthrich, K. The program XEASY for computer-supported NMR spectral-analysis of biological macromolecules. *J. Biomol. NMR* **1995**, 6, 1–10.

(61) Goddard, T. D.; Kneller, D. SPARKY, 3 ed.; University of California, San Francisco, CA, 2001.

(62) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 66, 125–132.

(63) The CCP4 suite: programs for protein crystallography. Collaborative Computational Project, Number 4. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, 50, 760–763.10.1107/S0907444994003112

(64) Blaszczyk, J.; Shi, G.; Yan, H.; Ji, X. Catalytic center assembly of HPPK as revealed by the crystal structure of a ternary complex at 1.25 Å resolution. *Structure* **2000**, 8, 1049–1058.

(65) Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, 60, 432–438.

(66) Vagin, A.; Steiner, R.; Lebedev, A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, 60, 2184–2195.

(67) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, 60, 2126–2132.

(68) Wlodek, S.; Skillman, A. G.; Nicholls, A. Automated ligand placement and refinement with a combined force field and shape potential. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, 62, 741–749.

(69) Joosten, R. P.; Womack, T.; Vriend, G.; Bricogne, G. Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, 65, 176–185.

(70) Maestro, version 10.3; Schrödinger, LLC: New York, NY, 2015.

(71) Schrödinger Suite 2015-3. *Protein Preparation Wizard*; *Epik*, version 3.3; Schrödinger, LLC: New York, NY, 2015. *Impact*, version 6.8; Schrödinger, LLC: New York, NY, 2015. *Prime*, version 4.1; Schrödinger, LLC: New York, NY, 2015.

(72) LigPrep, version 3.5; Schrödinger, LLC: New York, NY, 2015.

(73) West, R. T.; Garza, L. A., 2nd; Winchester, W. R.; Walmsely, J. A. Conformation, hydrogen bonding and aggregate formation of guanosine 5'-monophosphate and guanosine in dimethylsulfoxide. *Nucleic Acids Res.* **1994**, 22, 5128–5134.

Supporting Information

Structural Basis for the Selective Binding of Inhibitors to 6-Hydroxymethyl-7,8-dihydropterin Pyrophosphokinase from *Staphylococcus aureus* and *Escherichia coli*.

Matthew L. Dennis,^{#,†} Noel P. Pitcher,[#] Michael D. Lee,[#] Aaron J. DeBono,[#] Zhong-Chang Wang,^{#,‡} Jitendra R. Harjani,[#] Raphaël Rahmani,[#] Ben Cleary,[#] Thomas S. Peat,[†] Jonathan B. Baell,[#] and James D. Swarbrick^{*,#}.

[#]Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia.

[†]CSIRO Biosciences Program, Parkville, VIC 3052, Australia.

[‡]State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, People's Republic of China

Table of Contents

Table S1. X-ray crystallography data collection and refinement statistics for EcHPPK/AMPCPP/inhibitor complexes.

Table S2. X-ray crystallography data collection and refinement statistics for SaHPPK/AMPCPP/inhibitor complexes.

Table S3. Root-mean-square deviation for ligands using native and cross-docking in SaHPPK.

Table S4. Root-mean-square deviation for ligands using native and cross-docking in EcHPPK.

Figure S1. Surface plasmon resonance measurements of HPPK inhibitors.

Figure S2. Electron densities of compounds and cofactor analogues within their respective HPPK ternary complexes.

Figure S3. Chemical structures of reported compounds.

Table S1. Data collection and refinement statistics for EcHPPK/AMPCPP/inhibitor complexes

	5ETP (13)		5ETK (18)		5ETL (24)		5ETM (34)		5ETN (35)		5ETO (50)	
Space group	P 2 ₁		P 2 ₁		P 2 ₁		P 2 ₁		P 2 ₁		P 2 ₁	
Wavelength (Å)	0.9707		0.9707		0.9537		0.9537		0.9537		0.9537	
Unit-cell parameters (Å, °)	a = 35.95, b = 57.77, c = 38.64, α = 90, β = 116.11, γ = 90		a = 35.79, b = 57.69, c = 38.45, α = 90, β = 115.57, γ = 90		a = 69.45, b = 36.18, c = 121.33, α = 90, β = 90.32, γ = 90		a = 35.76, b = 57.63, c = 38.41, α = 90, β = 115.49, γ = 90		a = 35.88, b = 57.89, c = 38.61, α = 90, β = 115.12, γ = 90		a = 35.77, b = 57.75, c = 38.42, α = 90, β = 115.46, γ = 90	
Diffraction data^a												
Resolution range (Å)	34.70-1.05 (1.06-1.05)		34.68-1.09 (1.11-1.09)		45.81-1.82 (1.86-1.82)		34.67-1.46 (1.48-1.46)		34.96-1.40 (1.42-1.40)		34.69-1.07 (1.09-1.07)	
No. of unique reflections	66,451 (3,193)		57,389 (2,728)		54,352 (3,034)		24,472 (1,219)		28,184 (1,404)		61,436 (2,967)	
No. of observed reflections	468568		414,088		392,519		173,388		195,895		433,431	
Completeness (%)	99.7 (99.2)		97.7 (95.0)		99.6 (95.7)		99.8 (99.9)		99.8 (99.9)		98.6 (96.8)	
Data redundancy	7.1 (6.5)		7.2 (6.7)		7.2 (7.0)		7.1 (6.5)		7.0 (7.0)		7.1 (6.7)	
Mean I/σ(I)	13.4 (5.1)		14.0 (3.1)		11.0 (2.6)		11.6 (2.7)		9.3 (2.6)		10.1 (2.0)	
R _{merge}	0.087 (0.298)		0.073 (0.563)		0.115 (0.710)		0.097 (0.662)		0.111 (0.678)		0.097 (0.819)	
R _{p.i.m.}	0.053 (0.194)		0.044 (0.354)		0.069 (0.426)		0.059 (0.419)		0.069 (0.424)		0.060 (0.521)	
Refinement												
R _{free} (%)	12.5		13.6		24.1		23.9		21.9		15.0	
R _{cyst} (%)	10.7		11.6		20.2		20.4		18.2		12.9	
Size of R _{free} set (%)	4.8		4.9		5.0		5.7		4.8		4.8	
Protein molecules in the asymmetric unit	1		1		4		1		1		1	
Inhibitor molecules	1		1		4		1		1		1	
Cofactor-analogue molecules	1		1		4		1		1		1	
Water molecules	281		159		528		88		123		151	
RMSD from ideal values												
Bond lengths (Å)	0.019		0.017		0.014		0.021		0.029		0.019	
Bond angles (°)	2.10		1.98		1.79		2.21		2.55		1.97	
Mean B factors (Å ²)	6.7		10.5		14.4		12.9		12.6		10.2	
Ramachandran plot												
Favored (%)	99.38		98.76		98.54		98.76		97.90		98.15	
Outliers (%)	0		0		0		0		0		0	

S3

Table S2. Data collection and refinement statistics for SaHPPK/AMPCPP/inhibitor complexes

	SETV (13)	5ETR (34)	SETS (35)	SETQ (40)	SETT (50)
Space group	P 6 ₁	P 2 ₁	P 2 ₁	P 2 ₁	P 2 ₁
Wavelength (Å)	0.9537	1.0329	1.0329	0.9707	1.0329
Unit-cell parameters (Å, °)	a = b = 84.25, c = 52.41, α = β = 90, γ = 120	a = 47.58, b = 68.11, c = 53.23, α = 90, β = 106.12, γ = 90	a = 47.79, b = 68.58, c = 52.99, α = 90, β = 105.94, γ = 90	a = 47.63, b = 68.20, c = 52.82, α = 90, β = 106.05, γ = 90	a = 47.55, b = 68.40, c = 53.12, α = 90, β = 105.94, γ = 90
Diffraction data^a					
Resolution range (Å)	42.57–1.72 (1.75–1.72)	45.71–1.32 (1.34–1.32)	40.90–1.95 (2.00–1.95)	50.76–1.96 (2.01–1.96)	40.93–1.55 (1.58–1.55)
No. of unique reflections	22,673 (1,210)	77,438 (3,932)	24,055 (1,689)	23,517 (1,623)	47,134 (2,309)
No. of observed reflections	457,516	571,491	169,081	163,741	356,013
Completeness (%)	100.0 (100.0)	100.0 (100.0)	99.8 (99.8)	100.0 (99.8)	99.3 (98.2)
Data redundancy	20.2 (20.1)	7.4 (7.0)	7.0 (7.1)	7.0 (6.6)	7.6 (7.6)
Mean I/σ(I)	20.3 (5.0)	12.8 (2.2)	10.0 (2.9)	9.1 (3.2)	11.6 (2.4)
R _{merge}	0.124 (0.959)	0.077 (0.832)	0.149 (0.758)	0.117 (0.497)	0.099 (0.856)
R _{p,i.m.}	0.041 (0.314)	0.046 (0.521)	0.091 (0.464)	0.072 (0.319)	0.058 (0.504)
Refinement					
R _{free} (%)	20.7	15.2	21.8	21.6	19.2
R _{cryst} (%)	18.1	12.4	18.9	18.2	15.8
Size of R _{free} set (%)	5.0	5.0	5.1	4.9	5.1
Protein molecules in the asymmetric unit	1	2	2	2	2
Inhibitor molecules	1	2	2	2	2
Cofactor-analogue molecules	1	2	2	2	2
Water molecules	101	328	154	136	256
RMSD from ideal values					
Bond lengths (Å)	0.013	0.017	0.013	0.012	0.023
Bond angles (°)	1.85	1.94	1.75	1.88	2.21
Mean B factors (Å ²)	12.8	14.5	10.6	12.4	15.9
Ramachandran plot					
Favored (%)	98.09	98.45	98.71	98.06	98.41
Outliers (%)	0	0	0	0	0

Table S3. Root-mean-square deviation for ligands using native and cross-docking in SaHPPK.

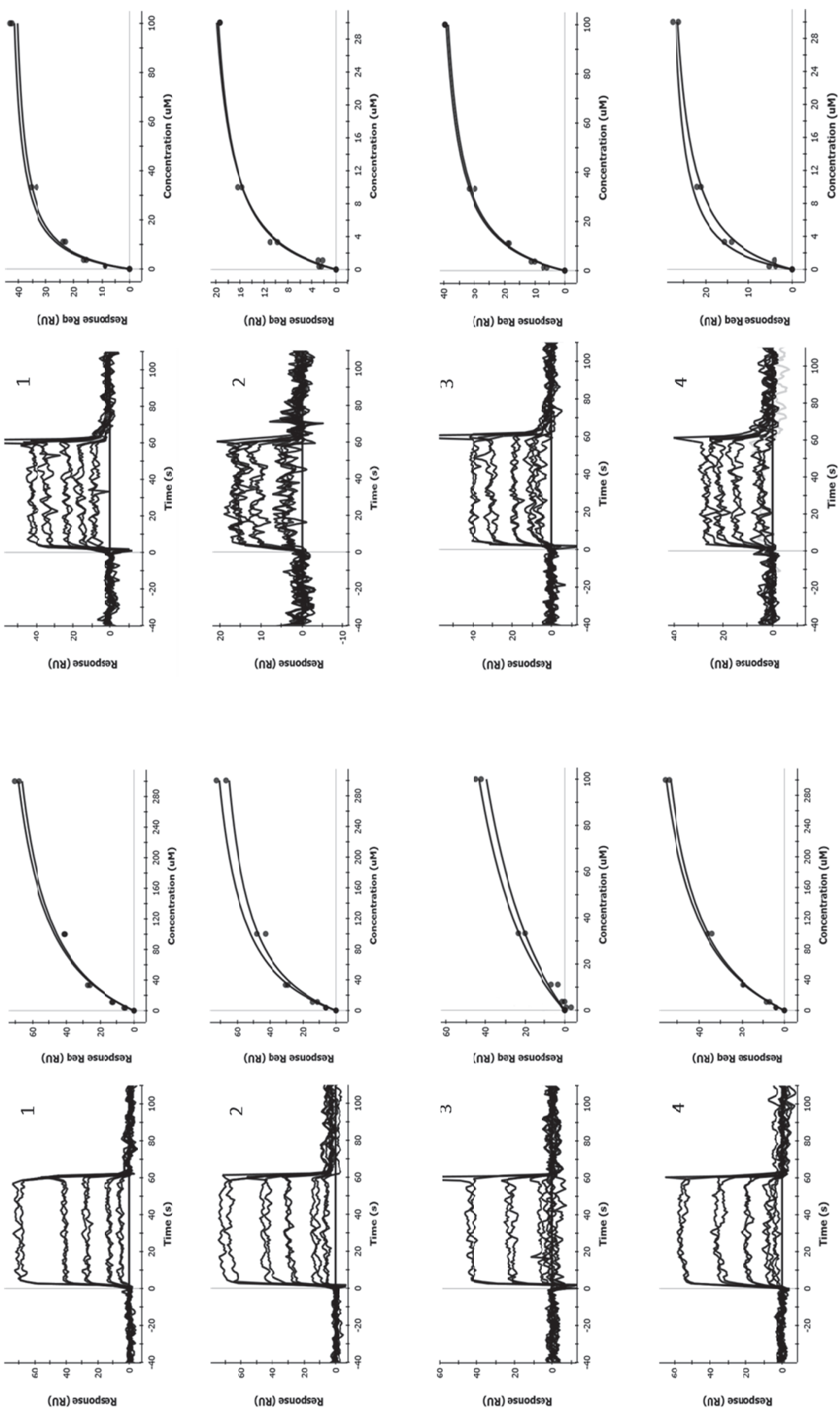
Compound	RMSD, native ligand docking (Å)		RMSD, cross-docking into 5ETR (Å) ^a
	SP mode	XP mode	
13	2.07	3.25	4.29
34	0.23	0.21	-
35	0.20	0.19	0.15
40	0.24	0.26	0.40
50	0.23	0.22	0.32

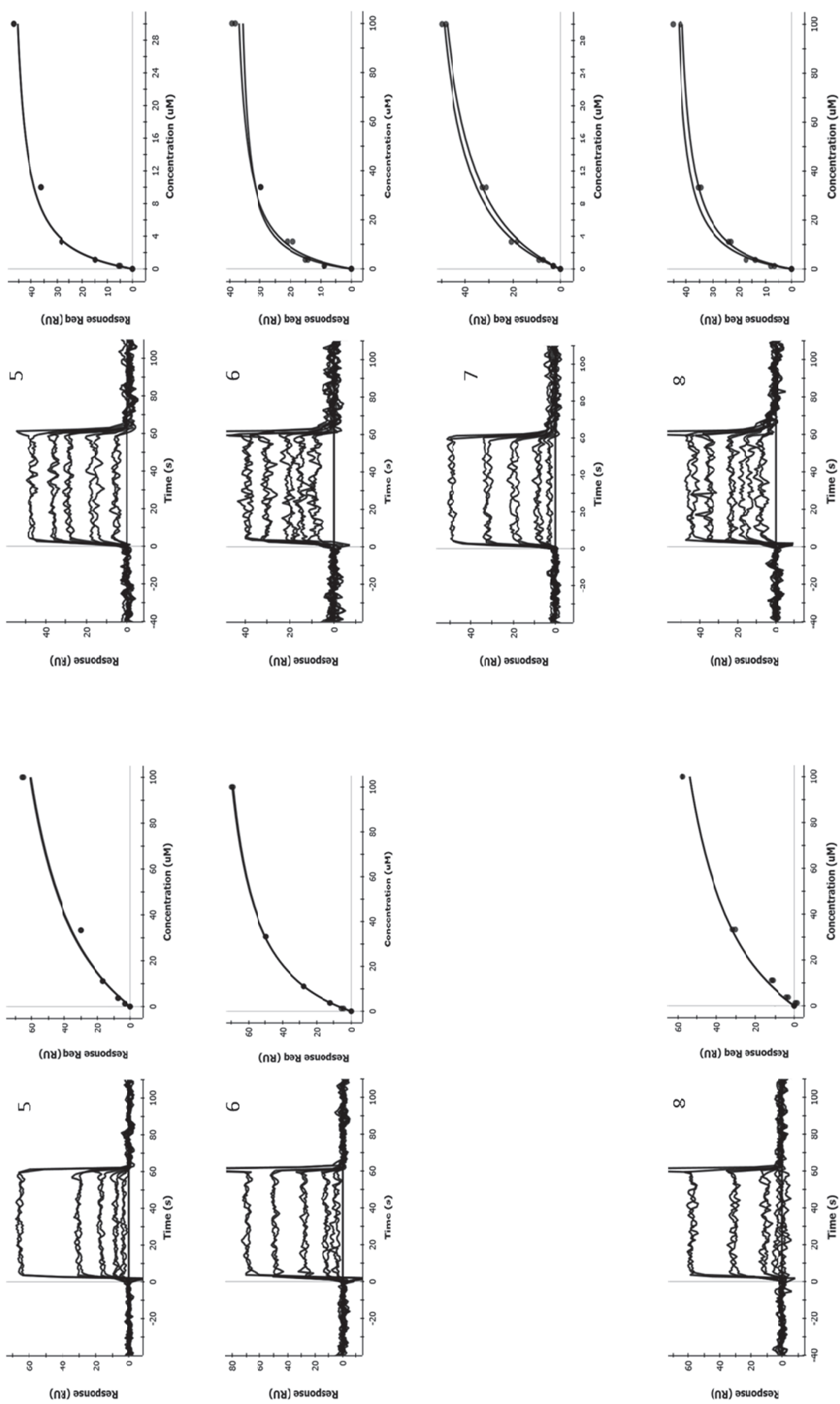
^aPerformed using Glide XP mode.

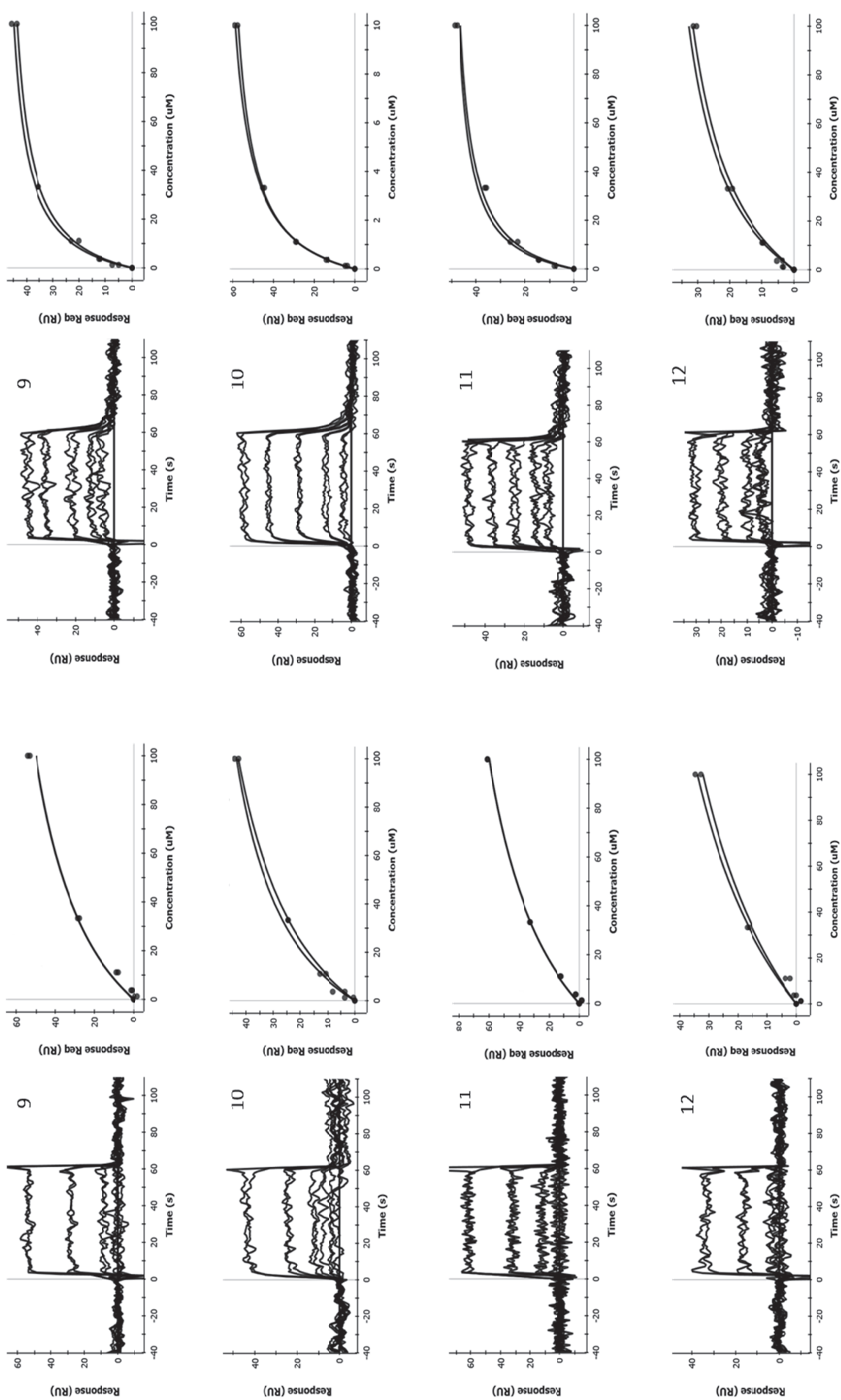
Table S4. Root-mean-square deviation for ligands using native and cross-docking in EcHPPK.

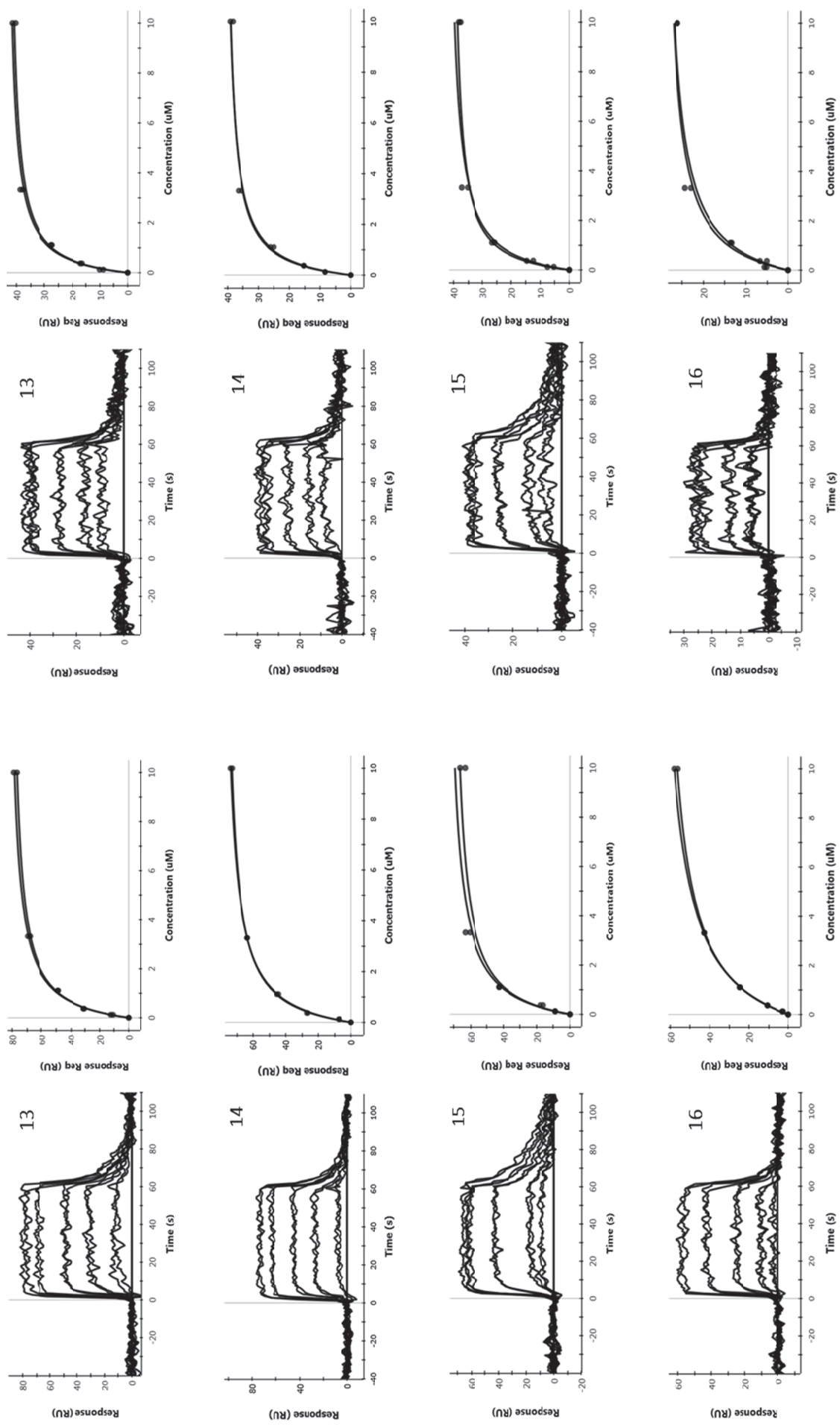
Compound	RMSD, native ligand docking (Å)		RMSD, cross-docking into 5ETP (Å) ^a
	SP mode	XP mode	
13	0.49	1.13	-
18	1.63	0.21	1.62
24	2.22	0.63	1.38
34	0.73	0.32	0.52
35	1.17	1.00	0.50
50	1.88	1.58	1.86

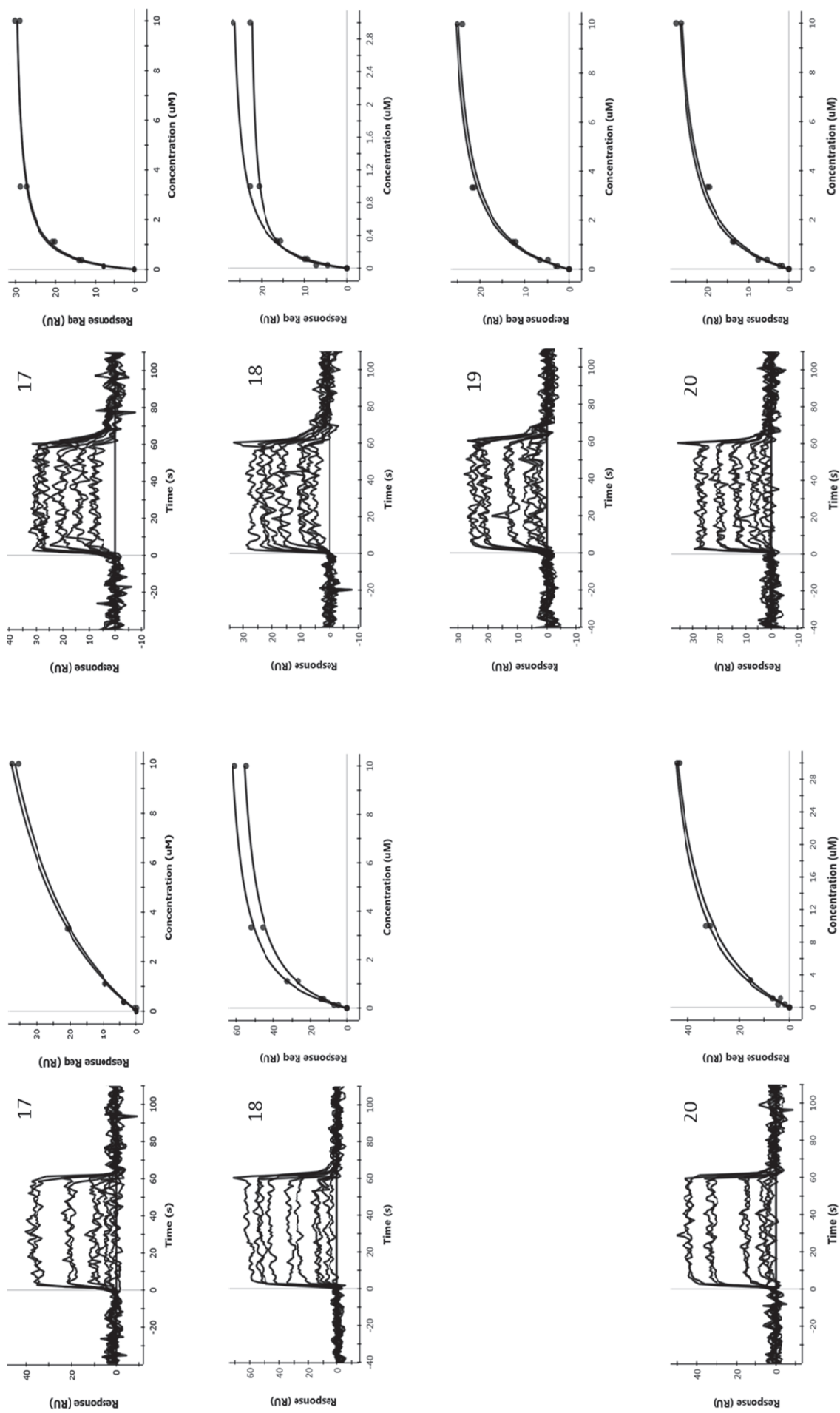
^aPerformed using Glide XP mode.



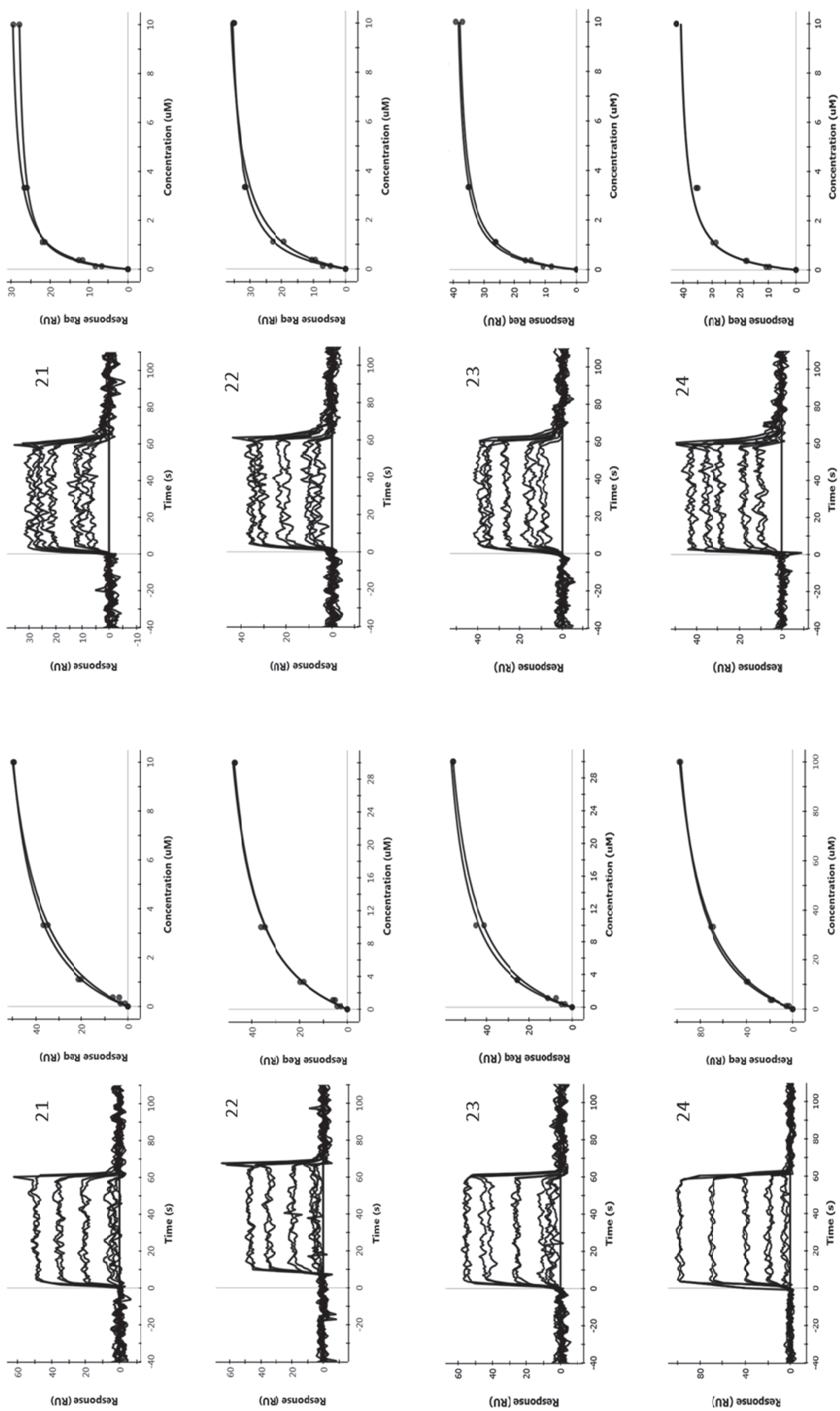


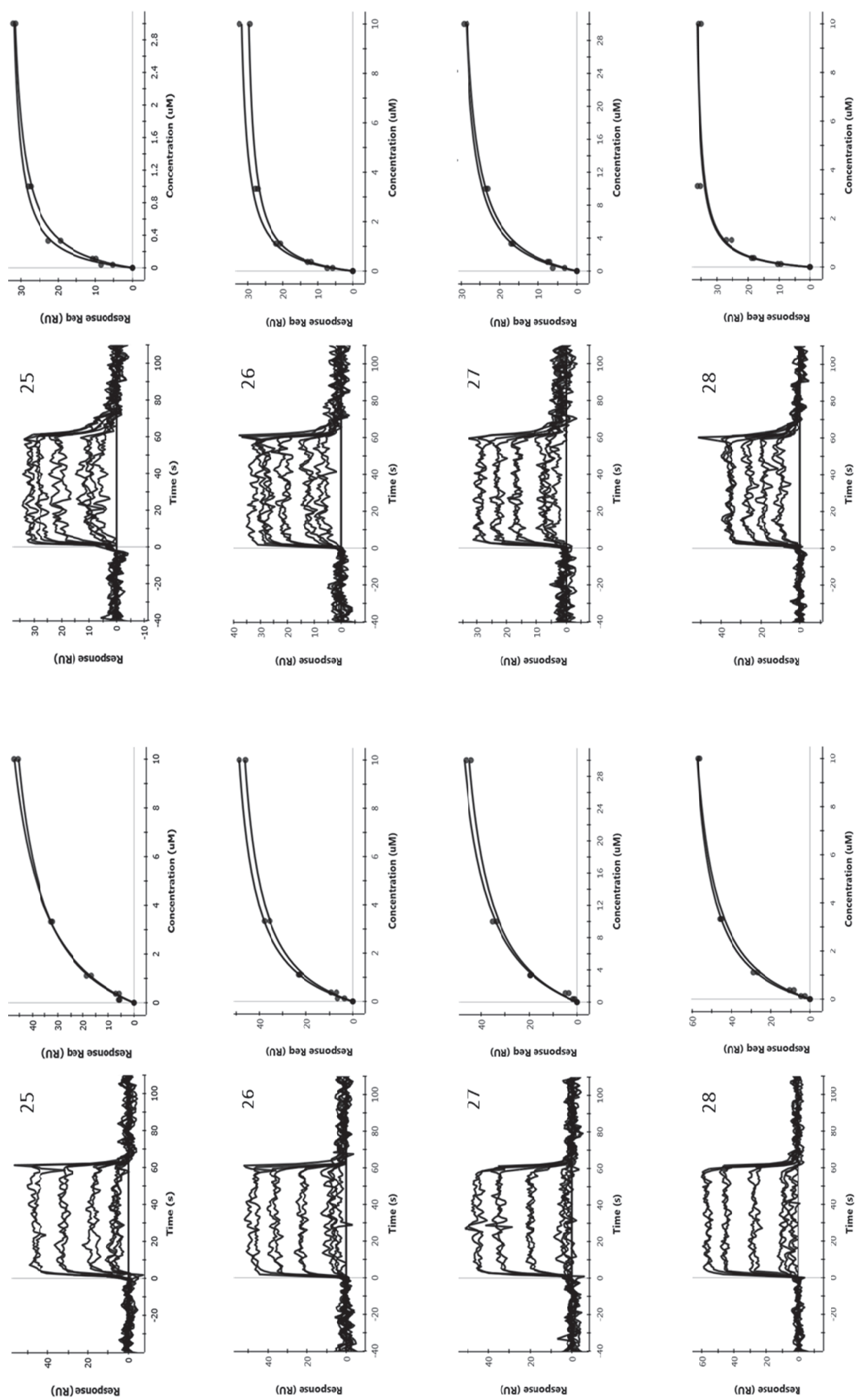




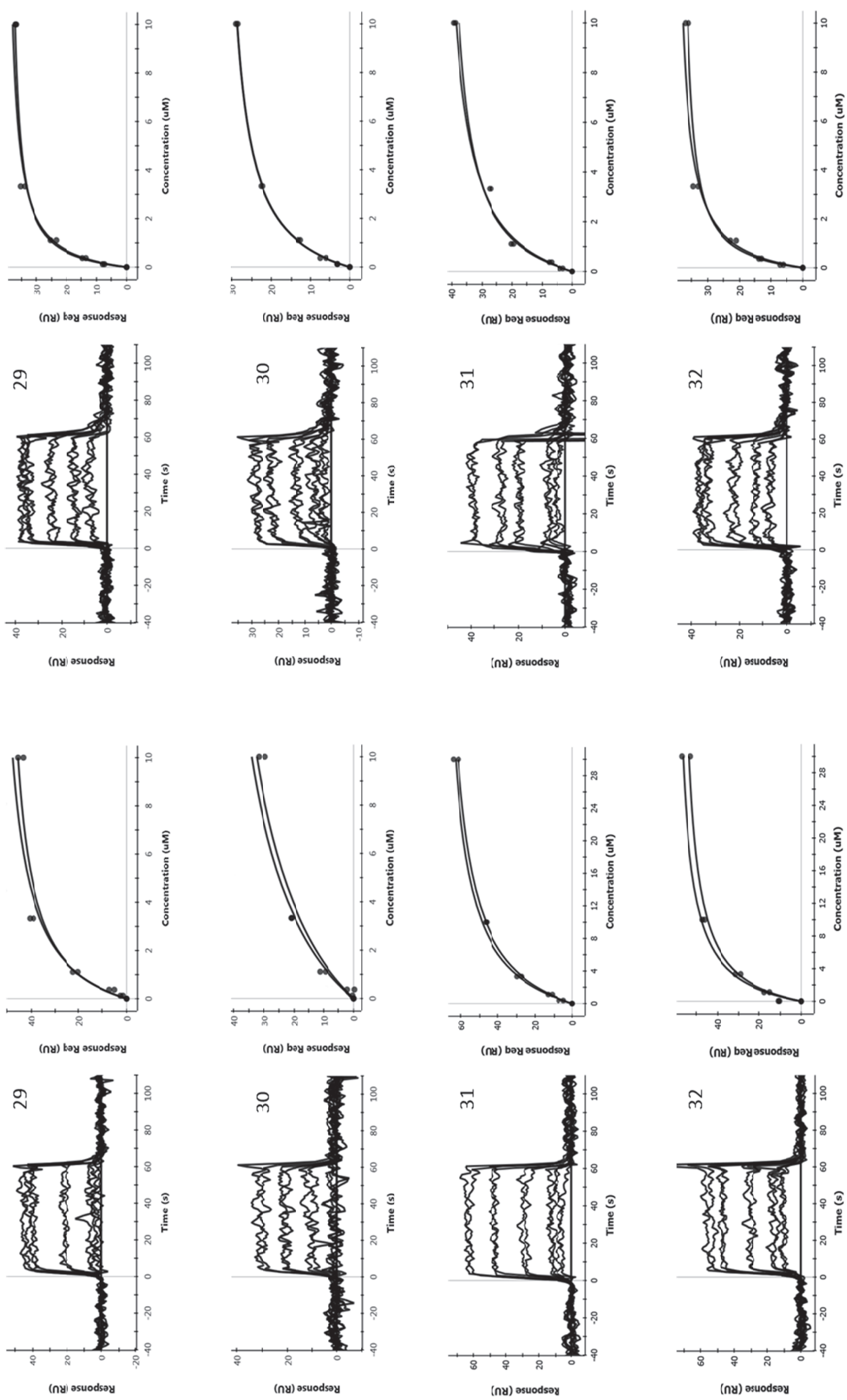


S10

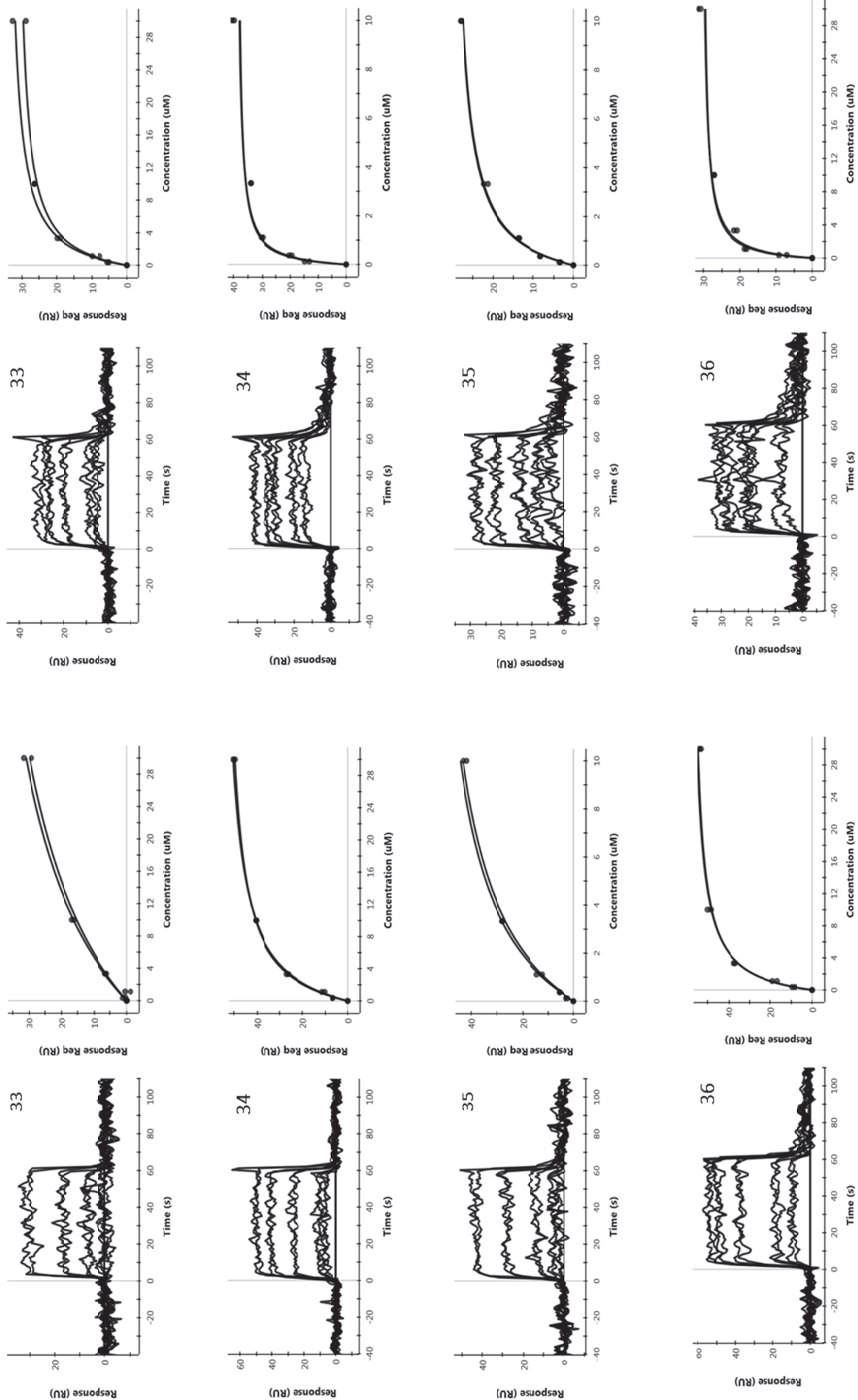


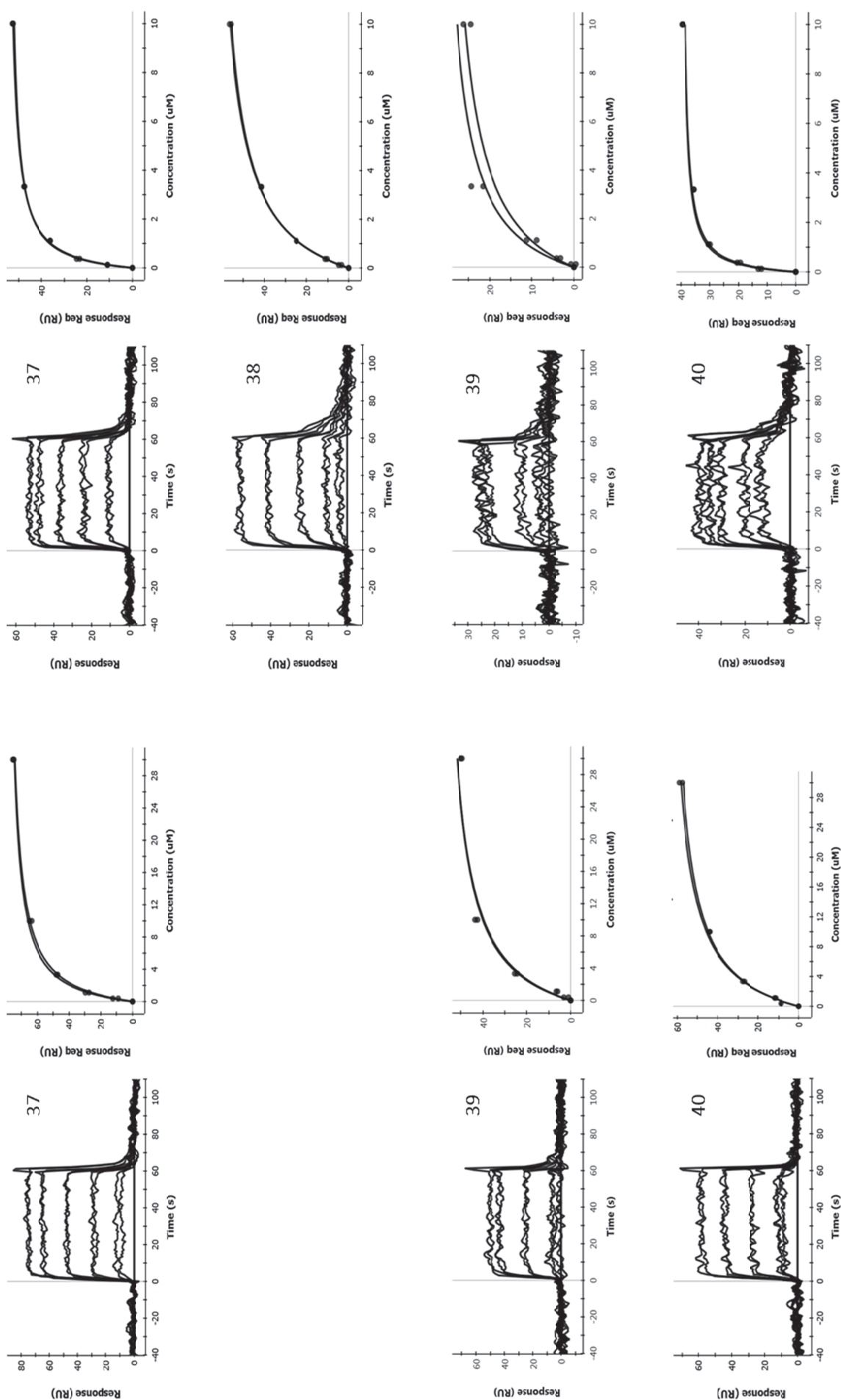


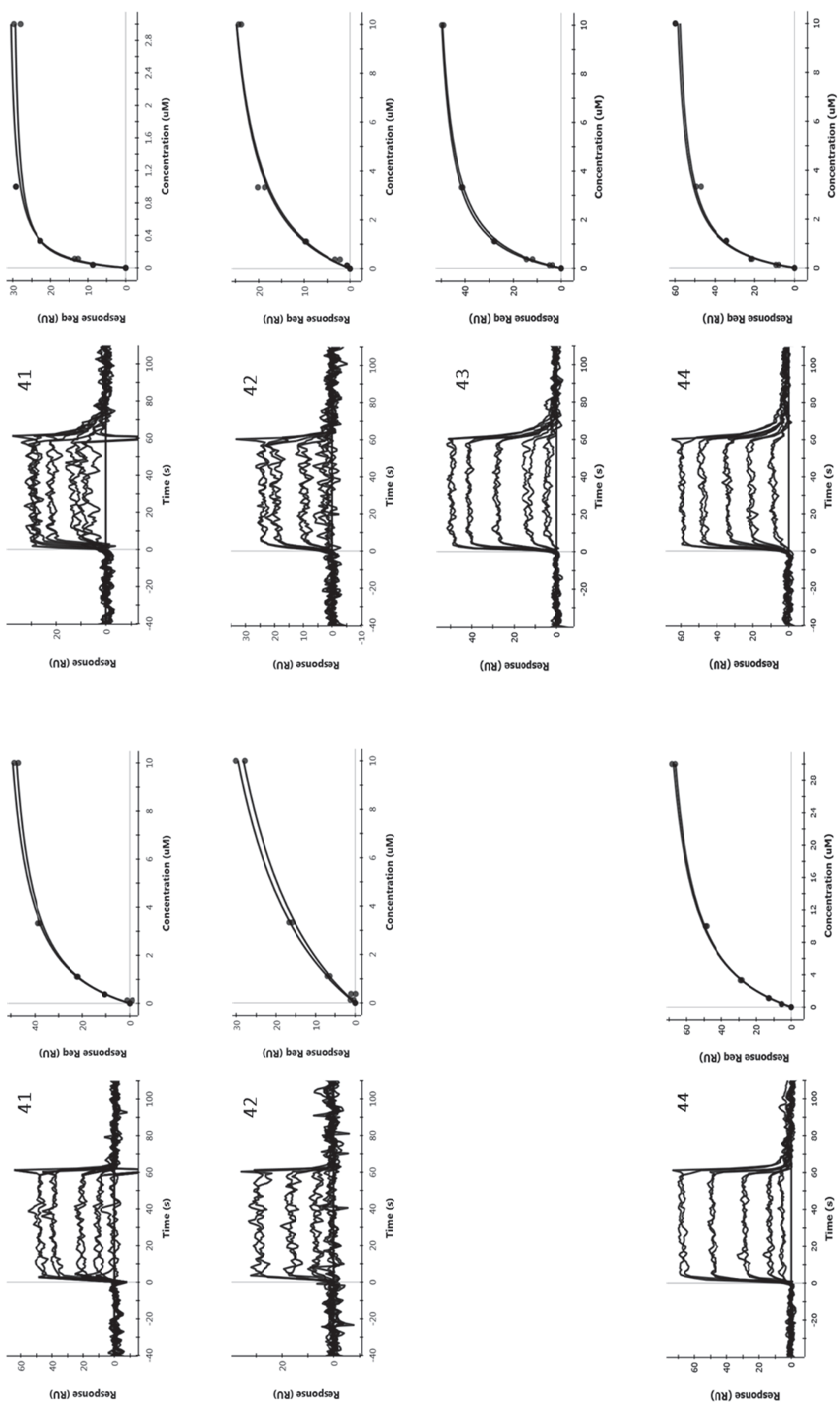
S12



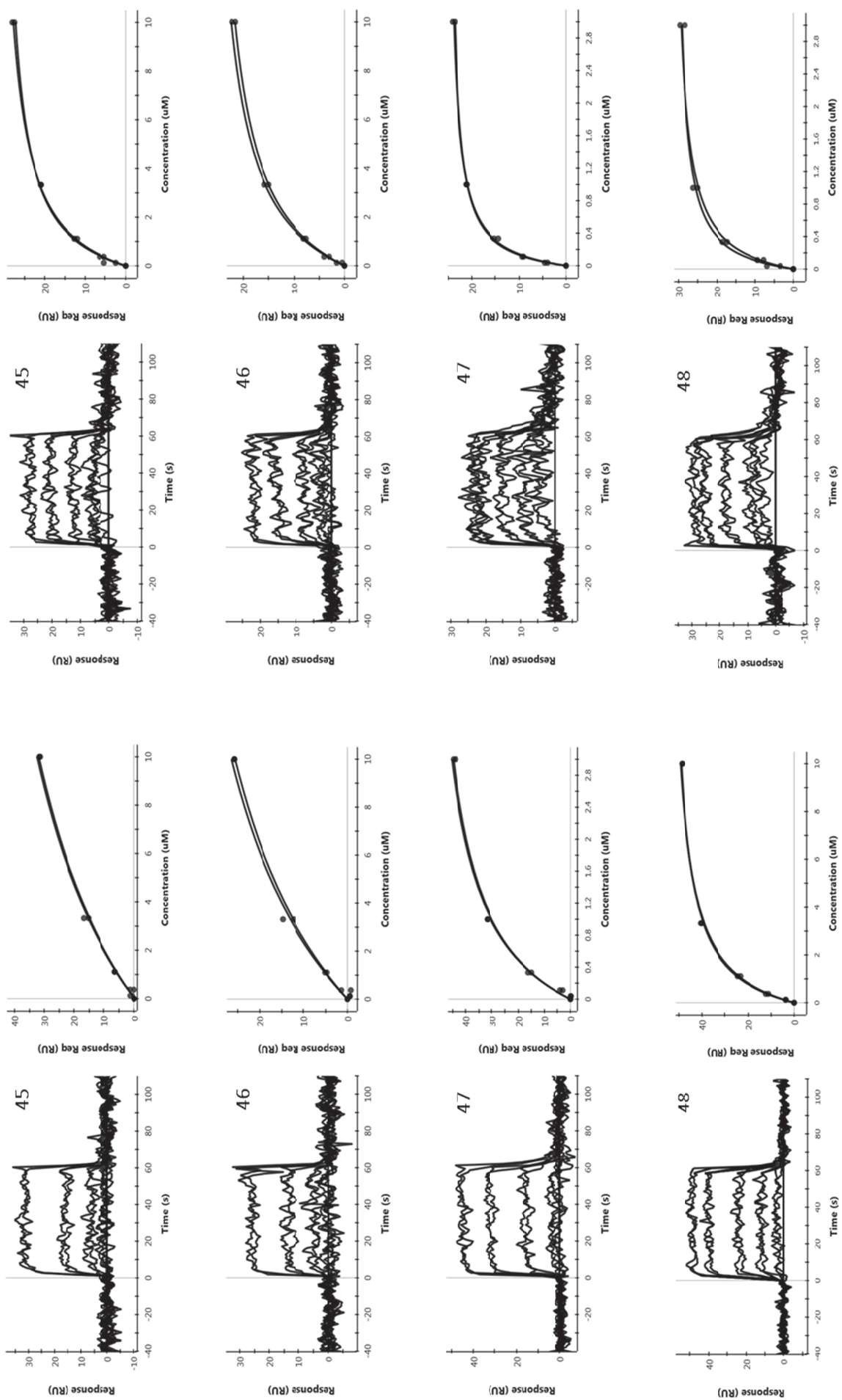
S13



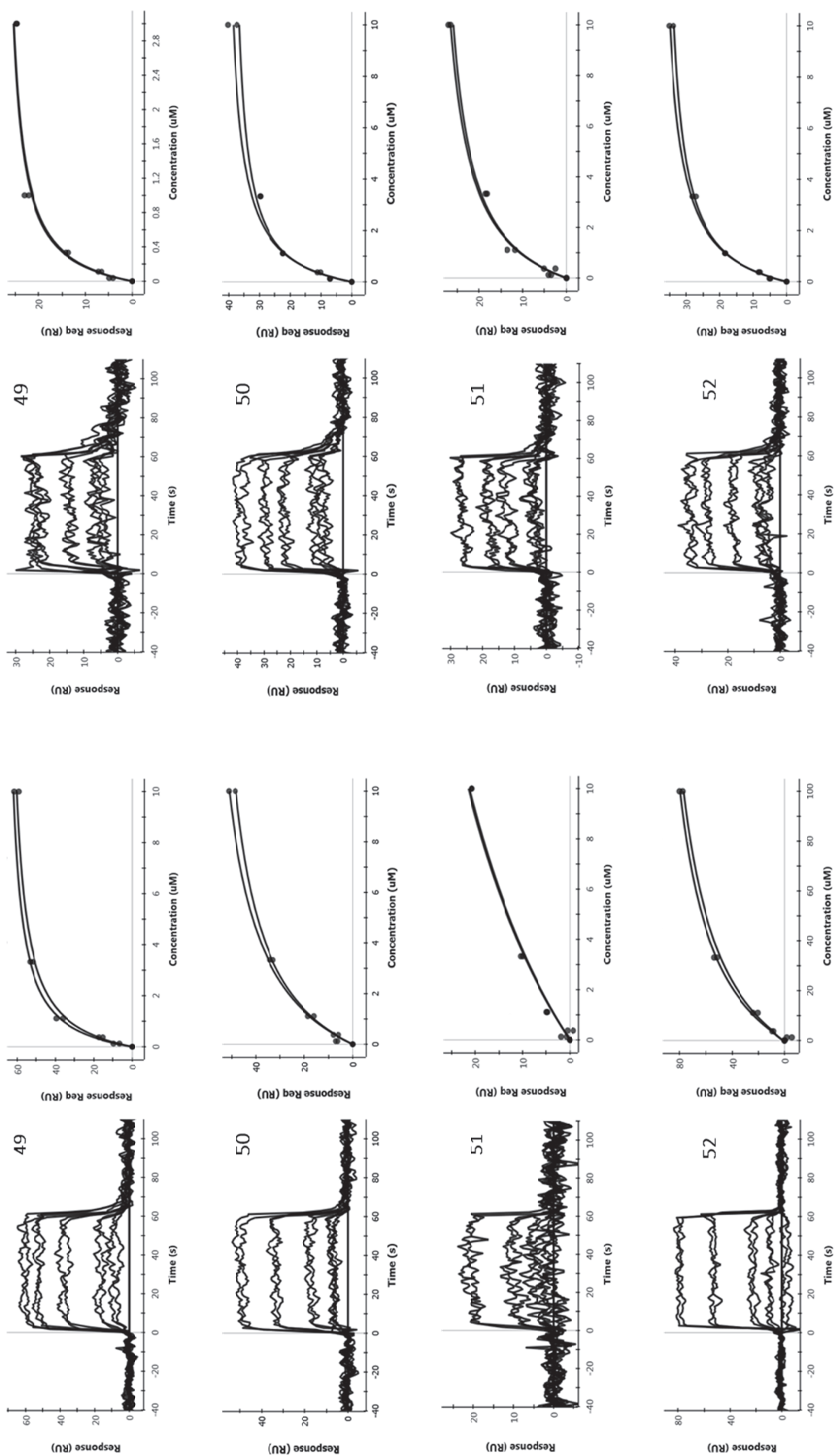




S16



S17



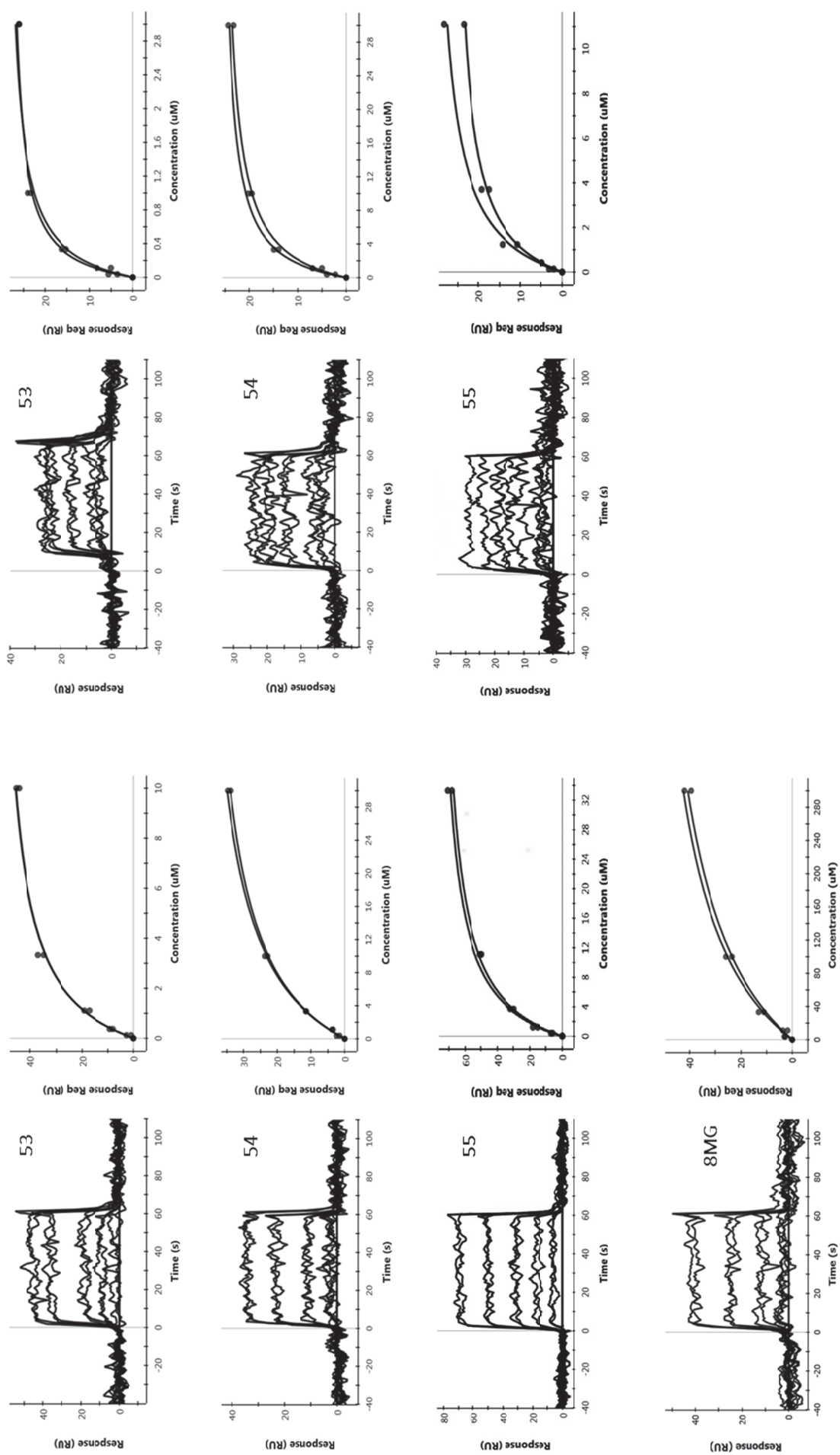


Figure S1. Surface plasmon resonance measurements of HPPK inhibitory compounds **1 – 55** and 8MG. Within each panel, the left sensorgram is testing against EcHPPK, and the right sensorgram is testing against SaHPPK. Compounds were tested in the presence of 1 mM ATP. The equilibrium dissociation constant (K_D) value for each compound was determined by equilibrium affinity analysis. The binding isotherms are shown to the right of the sensorgrams.

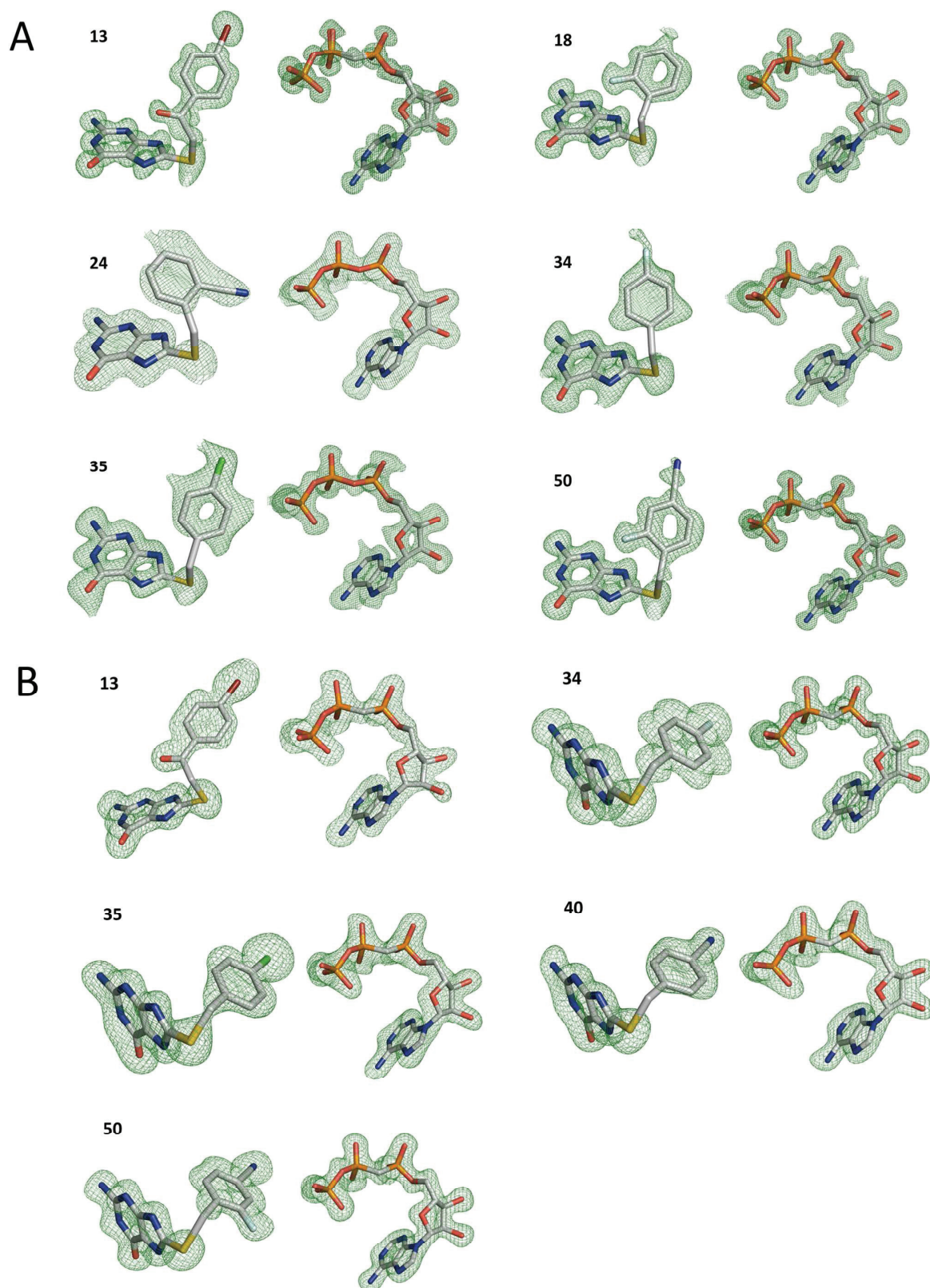
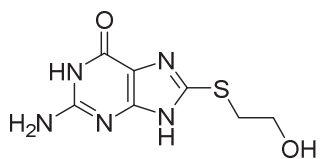
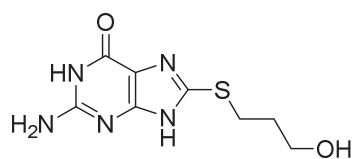


Figure S2. Electron densities of compounds and the corresponding complexed cofactor analogue, AMPCPP, bound to (A) EchPPK and (B) SaHPPK. $F_o - F_c$ difference density maps are contoured at 3.0σ .

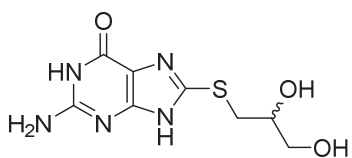
S21



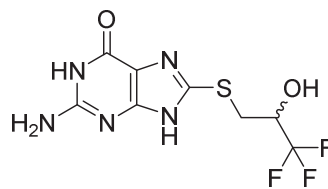
1



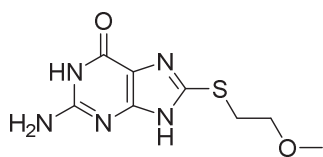
2



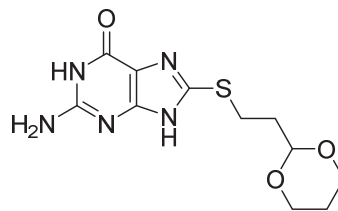
3



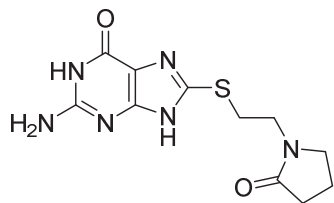
4



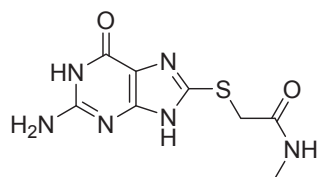
5



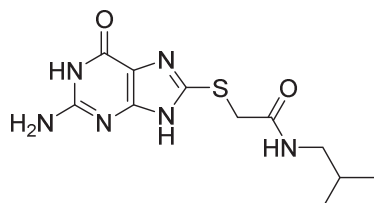
6



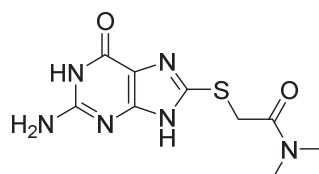
7



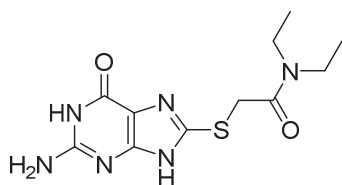
8



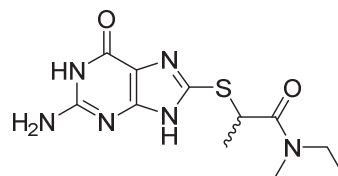
9



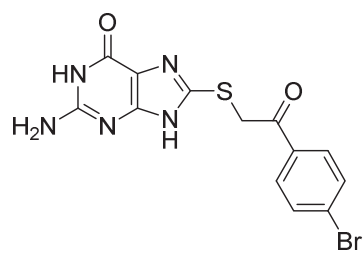
10



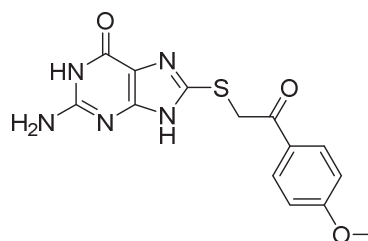
11



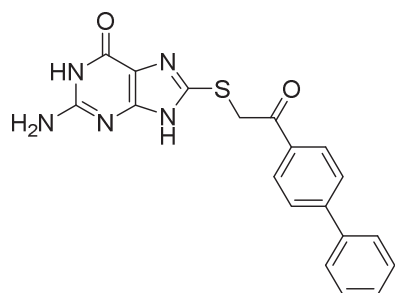
12



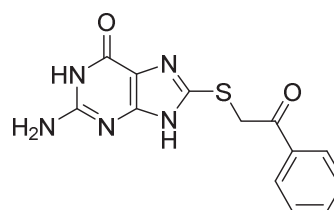
13



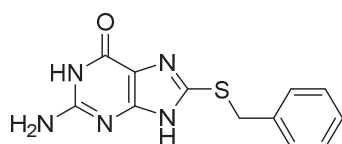
14



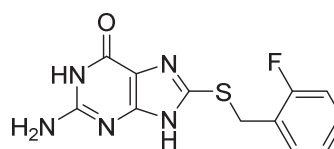
15



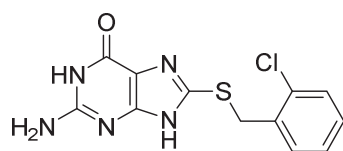
16



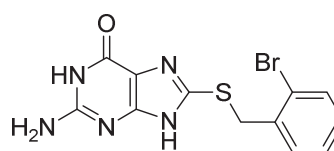
17



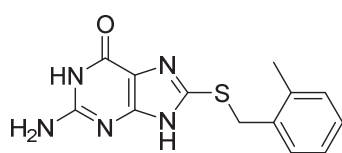
18



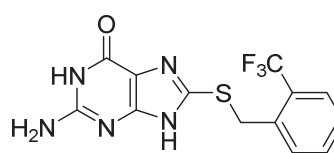
19



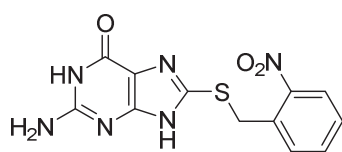
20



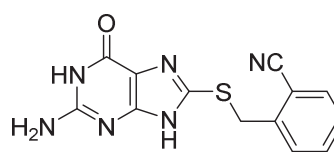
21



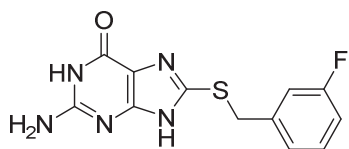
22



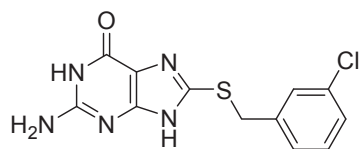
23



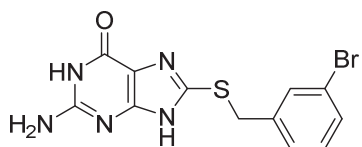
24



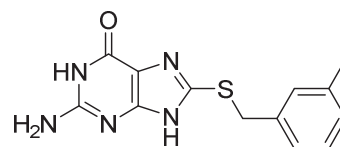
25



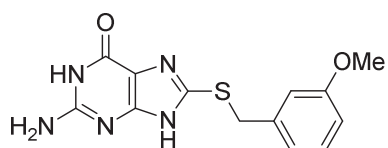
26



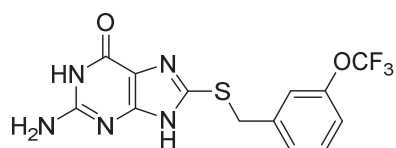
27



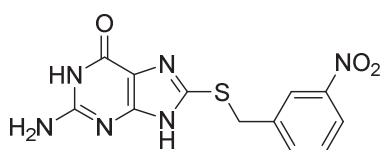
28



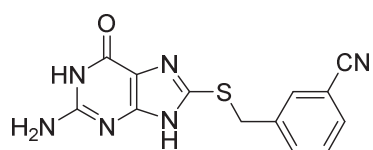
29



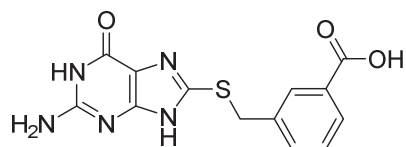
30



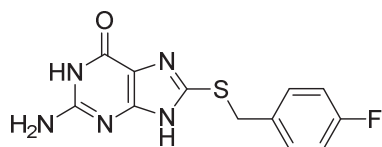
31



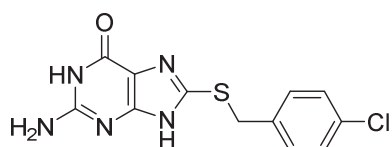
32



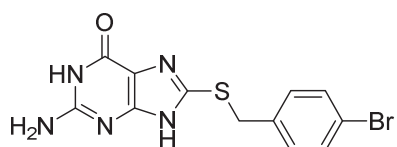
33



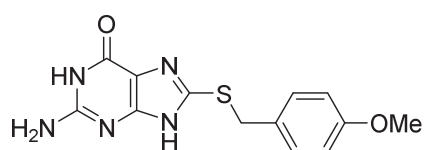
34



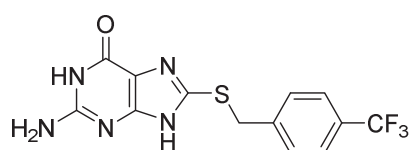
35



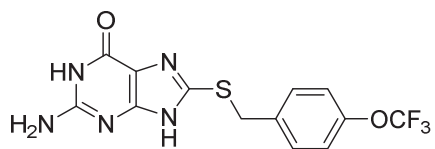
36



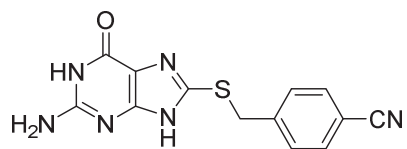
37



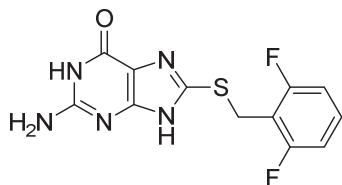
38



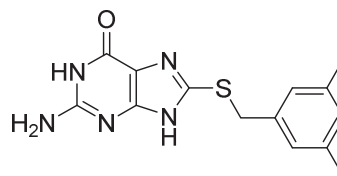
39



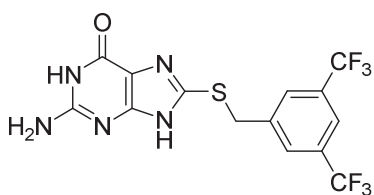
40



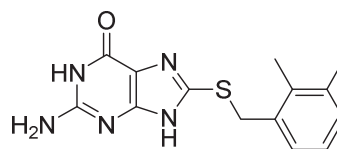
41



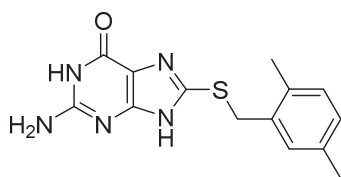
42



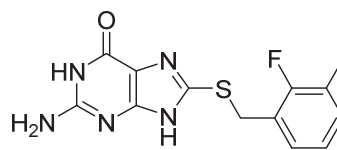
43



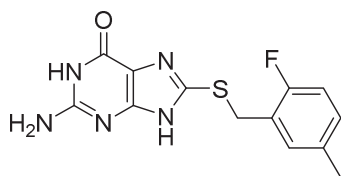
44



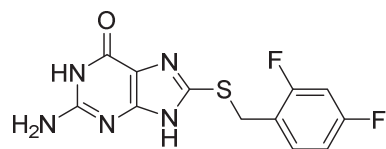
45



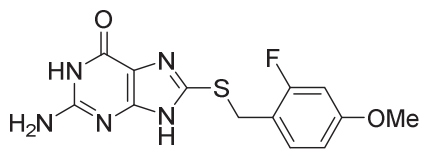
46



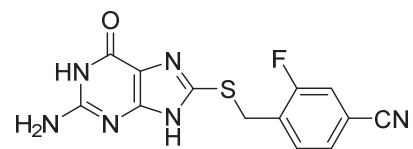
47



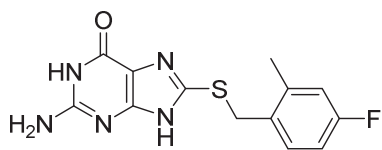
48



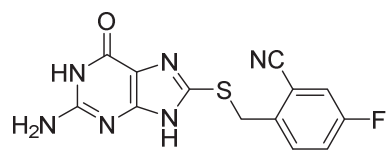
49



50



51



52

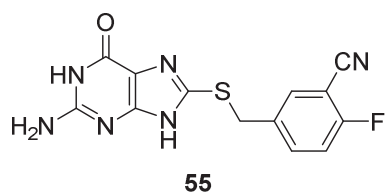
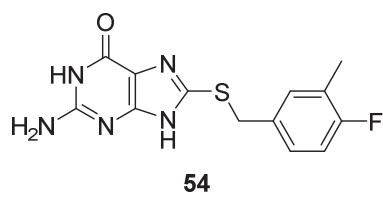
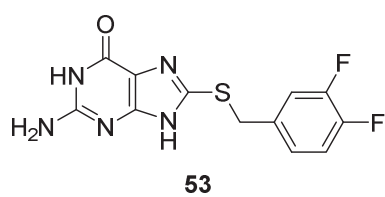
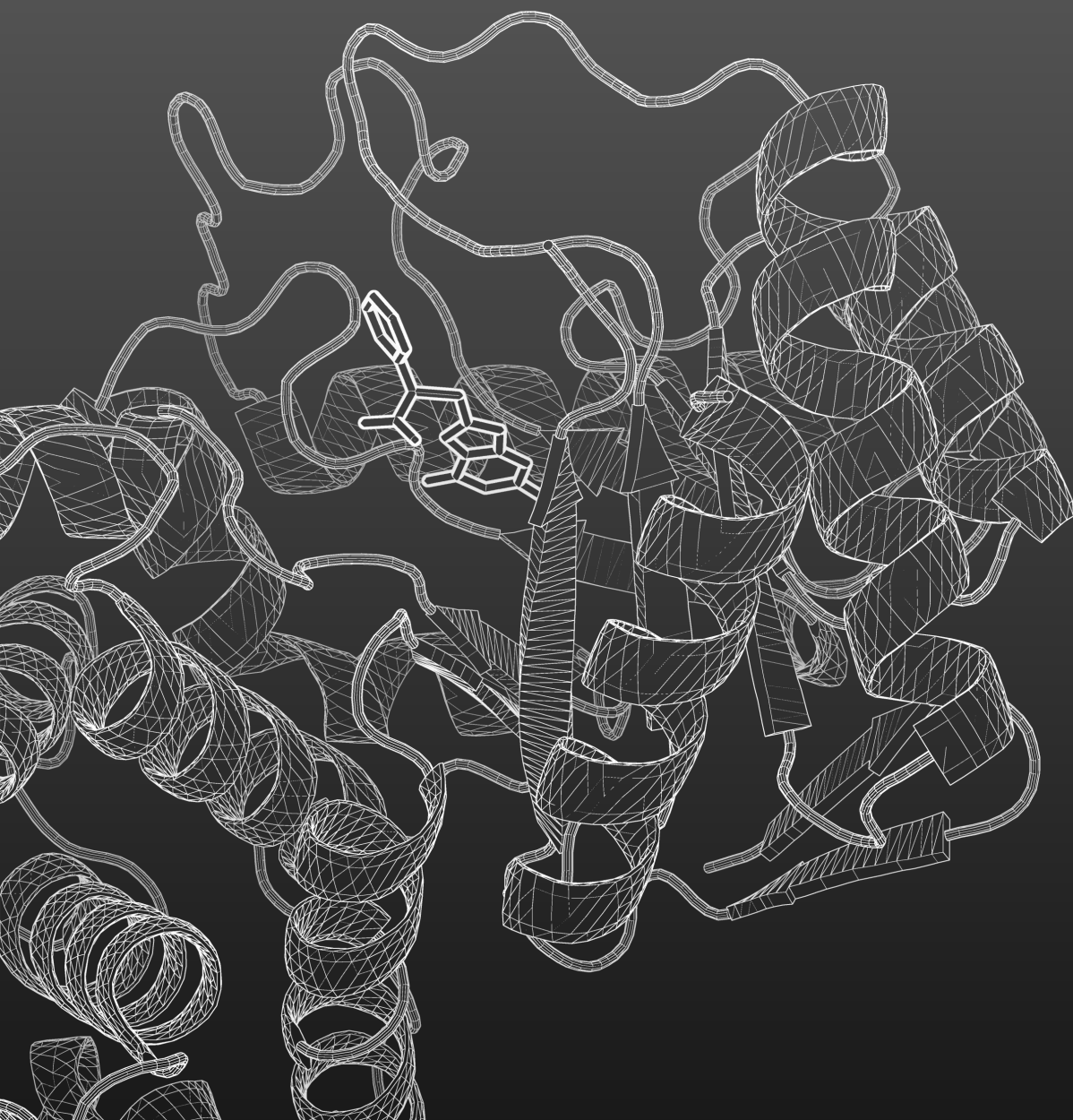


Figure S3. Chemical structures of reported compounds.

"8-Mercaptoguanine derivatives
as inhibitors of dihydropteroate
synthase from
Escherichia coli"

4



Presented in both Chapter 2 and the following paper, various compounds were discovered to bind to EcDHPS with low micromolar affinity. We were able to improve upon this affinity using the determined SAR, with compounds **45** and **46** (Figure 4.1) binding ~100-fold tighter than the parent, 8MG. For some time, however, a structural understanding of ligand binding eluded us — attempts to crystallise EcDHPS in complex with ligands of interest was proving problematic and the focus of the project shifted towards the more successful HPPK work, the EcDHPS work continuing in the background. This latter work comprised optimisation of the crystallants, protein concentration, drop size, temperature, and seeding used in crystallisation. Although robust crystal growth and diffraction was achieved, the diffraction was consistently in the range of 3.0–3.5 Å, with poor ligand density observed.

The concurrent use of NMR techniques with EcDHPS was similarly marred. The increased size of the DHPS protein in solution (~60 kDa dimer) versus HPPK (~20 kDa monomer) meant that deuterium labelling was now required.¹ *E. coli* cells possessing the EcDHPS plasmid were thus incrementally adapted to D₂O. Such D₂O-adapted cells are known to display significant variation in expression levels² and a process of colony selection was implemented,² whereby individual colonies from an agar plate were tested for expression levels; the highest expressing colony was then plated to repeat the process. Temperature, the duration of expression, and the concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG) added were optimised for use with high-cell-density expression methods.² Scaled-up expression returned significantly reduced protein levels compared to the equivalent H₂O expression, which is not unexpected, with ~500 mL of D₂O required per NMR sample (at a concentration suitable for HSQC-type experiments). While the ¹⁵N TROSY experiment obtained from a ~500 mL expression was very encouraging (Figure 4.2), the quantity of D₂O necessary for 3D NMR experiments for assignments was deemed prohibitively excessive. Furthermore, from a peak count, 50–60 amides were broadened and not observed, most likely due to μs motion of the loops and/or very slow back exchanging amides within the core. Subsequent structural efforts therefore concentrated on the use of crystallographic methods.

It was hypothesised that improving the stability of EcDHPS could improve its crystallisability and a differential scanning fluorimetry (DSF) assay was employed in this regard. A broad screen of pH (5.0–9.0) and buffers,³ with both high and low salt concentrations, followed by an additive screen (96 different additives), revealed a buffer that markedly improved the “melt temperature” of EcDHPS. Fortunately, a broad crystallisation

screen with this new buffer proved successful. Impressive crystal growth was observed, and a ligand-soaked structure of **45**/EcDHPS was solved to 1.88 Å. Co-crystallisation trials with various ligands, using these conditions, were likewise successful. Like EcHPPK, these EcDHPS crystals can form without ligand present. The UV-absorptive nature of the ligands again allowed the use of UV fluorescence as an indicator of ligand occupancy.

Compounds were tested in an antibacterial assay, however no inhibition of cell growth was observed. This lack of efficacy was overcome with the use of membrane-disrupting polymyxin antibiotics. When either polymyxin B or polymyxin E (colistin) was present at a sub-MIC concentration, the 8MG-derivatives were observed to exert a synergistic antibacterial effect. This was an important result, demonstrating the potential of these compounds to be developed into antibiotics, though highlighting the real need to improve upon the physicochemical properties to achieve membrane permeability.

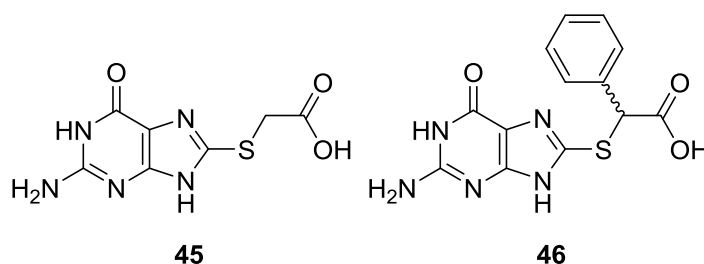


Figure 4.1. Chemical structures of compounds **45** and **46** reported in the following paper.

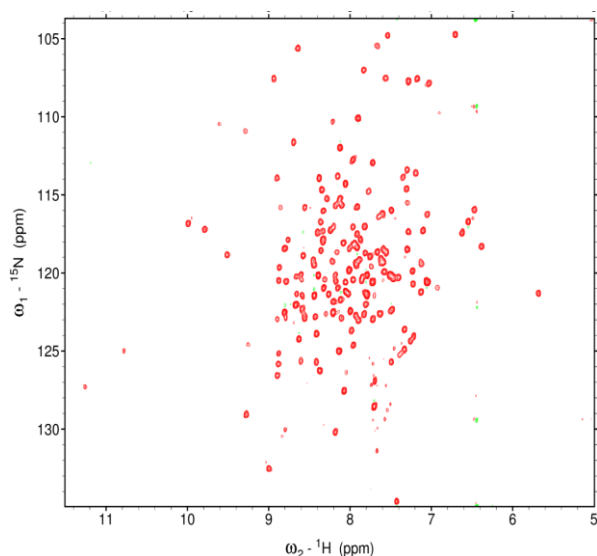


Figure 4.2 ^{15}N TROSY experiment recorded on a $\sim 80\ \mu\text{M}$ sample of $^{13}\text{C}/^{15}\text{N}/^2\text{H}$ DHPS from *E. coli* dissolved in 50mM HEPES, pH 7.5, and recorded for 5 h at 25 °C.

1. Tugarinov, V.; Hwang, P. M.; Kay, L. E. Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev. Biochem.* **2004**, *73*, 107–146.
2. Sivashanmugam, A.; Murray, V.; Cui, C.; Zhang, Y.; Wang, J.; Li, Q. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* **2009**, *18*, 936–948.
3. Seabrook, S. A.; Newman, J. High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Comb. Sci.* **2013**, *15*, 387–392.

Declaration for Thesis Chapter 4

Chapter 4 presents the prepared manuscript co-authored with Aaron DeBono, Michael D. Lee, Noel P. Pitcher, Zhong-Chang Wang, Sandeep Chhabra, Raphaël Rahmani, Olan Dolezal, Meghan Hattarki, Bim Graham, Thomas S. Peat, Jonathan B. Baell, and James D. Swarbrick.

Declaration by candidate

In the case of Chapter 3, the nature and extent of my contribution to the work was as the following:

Nature of Contribution	Extent of Contribution (%)
Protein expression and purification, NMR and X-ray crystallography experiments, SPR, DSF, biochemical assays, antibacterial assays, molecular modelling	70

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Aaron DeBono ¹	Synthetic chemistry, intellectual input	
Michael D. Lee ¹	Synthetic chemistry, intellectual input	5
Noel P. Pitcher ¹	Synthetic chemistry	5
Zhong-Chang Wang ³	Synthetic chemistry	
Sandeep Chhabra	Synthetic chemistry, SPR	
Jitendra Harjani ¹	Synthetic chemistry	
Raphaël Rahmani ¹	Supervision of chemists	

Olan Dolezal	SPR, intellectual input	
Meghan Hattarki	SPR	
Bim Graham	Intellectual input	
Thomas S. Peat ²	X-ray crystallography, intellectual input	
Jonathan B. Baell ¹	Synthetic chemistry, intellectual input	
James D. Swarbrick ¹	NMR experiments, intellectual input	

Declaration by co-authors

The undersigned hereby certify that:

- (1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors;
- (2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- (3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- (4) there are no other authors of the publication according to these criteria;
- (5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- (6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

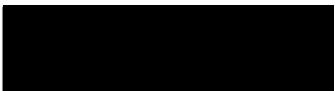
Location(s)

¹Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

²CSIRO Biosciences Program, Parkville, Victoria 3052, Australia

³State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, People's Republic of China

		Date
Signature 1		19/12/2016
Signature 2		
Signature 3		

		Date
Signature 2		19/12/16
Signature 3		19/12/16
Signature 4		20/12/16

8-Mercaptoguanine derivatives as inhibitors of dihydropteroate synthase from *Escherichia coli*

*Matthew L. Dennis,^{#†} Aaron J. DeBono,[#] Michael D. Lee,[#] Noel P. Pitcher,[#] Zhong-Chang Wang,^{#‡}
Sandeep Chhabra,^{#†} Raphaël Rahmani,[#] Olan Dolezal,[†] Meghan Hattarki,[†] Bim Graham,[#] Thomas
S. Peat,[†] Jonathan B. Baell,[#] and James D. Swarbrick.[#]*

[#]Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia.

[†]CSIRO Biomedical Program, Manufacturing, Parkville, VIC 3052, Australia.

[‡]State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093,
People's Republic of China

KEYWORDS

Folate biosynthesis, antimicrobials, X-ray, 8-mercaptopguanine, SAR, *E. coli*, DHPS.

ABSTRACT

Dihydropteroate synthase (DHPS) is an enzyme of the folate biosynthesis pathway that catalyzes the formation of 7,8-dihydropteroate (DHPT) from 6-hydroxymethyl-7,8-dihydropterin pyrophosphate (DHPPP) and *para*-aminobenzoic acid (pABA). DHPS has been the long-standing target of the sulfonamide class of antibiotics that compete with pABA, however resistance has emerged due to mutations in the flexible pABA region. Targeting the more rigid (and conserved) pterin site has been suggested as an attractive approach to mitigate the rise of similar mutant-based resistance strains and counter current sulfa drug resistance isolates. Following our related work targeting the pterin site of the adjacent enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), we present derivatives of 8-mercaptopguanine, a fragment that binds the pterin site in both enzymes, and quantify binding for *Escherichia coli* DHPS (EcDHPS). Ligand-bound EcDHPS crystal structures delineate the structure-activity relationship observed and provide structural insight for the development of novel pterin-site inhibitors with the potential for dual HPPK/DHPS inhibition.

INTRODUCTION

The sulfonamide class of antibiotics was discovered in the 1930s, and sulfonamides and related sulfones are still used today to treat infectious diseases including malaria, tuberculosis, *Pneumocystis jirovecii* pneumonia (the most common fungal infection in HIV and AIDS infected patients), and methicillin-resistant *Staphylococcus aureus*.¹⁻⁴ Sulfonamides target the folate-pathway enzyme dihydropteroate synthase (DHPS), which catalyzes the condensation of *p*-aminobenzoic acid (pABA)

and 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate (DHPPP) into dihydropteroate (DHPt) (Figure 1). Sulfonamides exert their effect through competing with pABA by reacting to form “dead-end” drug-pterin adducts, depleting cellular DHPPP as well as inhibiting downstream enzymes.⁵⁻⁹

Despite decades of success in the clinic, only recently through crystallization studies has the catalytic mechanism and the structural basis behind sulfa drug resistance been revealed.¹⁰ Similarly, the structural basis for off-target effects leading to the commonly observed sulfa allergy side effect has been established.¹¹ These structural details and other studies promise to help the development of newer generation sulfa drugs with improved therapeutic effect and longevity.^{12, 13} In many cases, sulfa drug resistance is a consequence of point mutations in the folP gene altering residues near to the pABA site.¹⁴ The pterin-binding site of DHPS has therefore been recognized as an alternative druggable pocket to reinvigorate the design of novel antimicrobials for this classic and validated drug target. Crystal structures from various organisms^{10, 15-24} reveal this to be a rigid pocket lined with highly conserved residues located deep in the enzyme, in stark contrast to the pABA site that is nestled in the flexible loop regions. Taking advantage of the rigid and conserved nature of the pterin pocket is potentially beneficial; under selective pressure, viable mutations conferring resistance to pterin-site inhibitors would be difficult to evolve without compromising activity. The conserved nature of the pterin pocket indicates that a suitably designed pterin-site agent could potentially illicit broad spectrum activity and also be effective against organisms currently exhibiting sulfa drug resistance.²⁵

Interest in pterin site inhibitors started in the 1970s when 4,5-dioxo-1,4,5,6-tetrahydropyrimido[4,5-*c*]pyridazines were developed to inhibit *Escherichia coli* DHPS (EcDHPS) by Burroughs Wellcome Co. (London, UK)^{26, 27} and a decade later various monocyclic 6-(alkylamino)-5-nitrosoisocytosines, including 6-methylamino-5-nitrosoisocytosine (MANIC, Figure 1), were published by Lever et al.^{28, 29} Both compound series were pterin-site inhibitors, as verified through recent crystallographic studies with *Bacillus anthracis* DHPS (BaDHPS) by the White group.^{17, 30} The same group also identified

other pterin-like compounds, such as 8-mercaptopguanine (8MG), through similarity searching of the pterin moiety.^{30, 31} Several of these were crystallized in the pterin site, despite not showing enzyme inhibition *in vitro*, as in the case of 8MG.³⁰

We have been generating focused libraries around the 8MG core since showing that the parent compound inhibits 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK, the enzyme preceding DHPS in the folate pathway), through binding at the pterin site.³²⁻³⁵ During our studies, we noticed that several of our 8MG derivatives also bound to EcDHPS,³⁴ which supports the design philosophy that pterin-like analogues may illicit beneficial polypharmacological activity.³⁶ In this study, we extend our work on the 8MG series and report crystal structures of several novel compounds in complex with EcDHPS, elucidating the binding mode and identifying key interactions for binding. Equilibrium binding constants were measured using SPR and, in two cases, sub-micromolar affinities were observed that constitute ~100-fold tighter binding than that of the parent compound, 8MG (**45**, $K_D = 830$ nM and **46**, $K_D = 650$ nM). Several of these new compounds displayed inhibition of *E. coli* K12 in a cell-based assay.

RESULTS AND DISCUSSION

Crystallization of EcDHPS. Initially, protein crystallization trials were performed using Tris-HCl (pH 8.5) as the protein buffer. However, poor reproducibility of crystal growth, suspected to arise from stability issues, led us to perform a differential scanning fluorimetry (DSF) assay³⁷ to screen for more suitable buffer conditions. A buffer solution containing 50 mM trisodium citrate, 150 mM NaH_2PO_4 , and 350 mM K_2HPO_4 , pH 6.9, was thus determined to provide EcDHPS with markedly improved stability (Figure S5), at both low and high temperatures, over the previously reported conditions,^{15, 38} which enabled consistent crystal growth and good diffraction for a range of complexes.

C⁸- and N⁹-functionalization of 8MG. Previously, 8MG was crystallized in BaDHPS, despite showing no inhibition (at 250 μ M) in a radiolabelled pABA assay during a pterin-like compound screen.³⁰ 8MG is a small fragment with respectable ligand efficiency for EcDHPS binding (K_D = 76 μ M over 12 heavy atoms³⁴). It represents a close scaffold hop from the pterin core, sharing the same isocytosine ring and a synthetically useful mercapto group through which to probe extension from the 8 position. Elaboration in this direction is justified by the fact that the C⁸-S⁸ vector points out from the deep pterin pocket in DHPS but is also oriented $\sim 30^\circ$ away from the methylene linker of the native substrate. While direct derivatization of the pterin scaffold has been carried out in earlier work (vide supra), elaboration of 8MG has not been investigated as a means to generate pterin-site inhibitors of DHPS. Our current study now continues with a structure-based investigation of a series of 8MG derivatives as potential EcDHPS inhibitors.

Both the N⁷ and N⁹ positions looked to be rather poor choices for elaboration of the 8MG core from simple superposition of the 8MG on the pterin moiety in the EcDHPS structure. For N⁹, the equivalent nitrogen group of the pterin substrate participates in a hydrogen bond with residue D96.¹⁵ Methylation at this position (and therefore removal of this bond), as per MANIC and other inhibitors,^{17, 30, 31} is tolerated, however there appears limited space to extend this substituent further. To test this, three simple 8MG derivatives with either ethyl (**1**), benzyl (**2**), or phenethyl (**3**) N⁹-functionalization were tested by SPR (Table 1). Indeed, these compounds all had reduced affinity for EcDHPS compared to 8MG, returning K_D values of 420 μ M, 350 μ M, and 160 μ M, respectively. The crystal structures of the S⁸-functionalized analogues reported in this study (vide infra) show this may result from the addition of unfavorably close contacts to the D96 residue. The affinity trend for these compounds appears to indicate, however, that as the N⁹-substituent increases in size, additional interactions to the protein are formed that begin to compensate for these effects. Inspection of the EcDHPS structure reveals that extension from the N⁷ position is similarly predicted to be highly unfavorable, since the N⁷ atom of 8MG is part of a bifurcated hydrogen bonding network involving the sidechain of residue K221 and the carbonyl of 8MG.

Guanine analogues with various functionalities at the C⁸ position — bromine (**4**), methyl (**5**), hydroxyl (**6**), and morpholine (**8**) — were tested (Table 1). Of these, only **4** showed improved binding relative to 8MG (~2-fold), returning an affinity of 42 μ M. N⁹-methylation of **6** improved binding 2-fold (**7**, K_D = 101 μ M); apart from the direct interactions from the introduced substituent, N⁹-methylation potentially improves binding by more favorable desolvation as well as promoting the bound tautomer, i.e. an unsaturated N⁷-C⁸ bond, with the resulting HBA at N⁷ interacting with K221. Considering that extension from N⁹ showed limited potential, strategically a thiol represented the obvious group for elaboration, offering synthetic ease for a range of chemical extensions. We therefore focused on functionalization at the S⁸ position.

S⁸-benzyl derivatization results in “flat” SAR. Previously we demonstrated that derivatization from the S⁸ position improves the affinity of 8MG for EcDHPS, with benzyl functionalization yielding binding constants as low as 8.5 μ M (compound **9**).³⁴ Here, further substitution from the benzyl ring was investigated. Testing of various 8MG-benzyl derivatives, with small substitutions at all positions (ortho, meta, para), in addition to naphthalene (**28**, **29**), phenylethynyl (**30**), and pyridine (**31**) analogues, revealed a particularly flat SAR (Tables 2 and 3), with affinities in the range of 7–16 μ M for 21 of the 22 compounds tested. The 8MG-benzyl derivative possessing *o*-chloro substitution (**10**) bound the tightest of these compounds, with an affinity of 7.0 μ M.

A crystal structure of EcDHPS in complex with compound **9**, featuring an *o*-fluorobenzyl group, was solved to 1.84 Å resolution. Two protomers are present in the asymmetric unit, contrasting with the one protomer observed in the three previously published EcDHPS crystal structures,¹⁵ where the dimer interface was present through crystallographic symmetry. The new complex is similar, with overall backbone rmsd values of 0.46–0.54 Å, when compared to these three previous EcDHPS structures (PDB ID: 1AJ2, 1AJ0, and 1AJZ).¹⁵ The biggest differences between the two monomers of the **9**/EcDHPS complex are for residues 57–70 (loop 2) and 143–151 (loop 5, absent in monomer B). Loop 1 (residues 25–34) is also disordered for both monomers. A key difference between the

pterin-binding sites of the monomers is found in the orientation of catalytic residue D96. This interacts with the ligand in monomer A, but is instead oriented towards residue D56 in monomer B — a trend consistent for all the crystal structures in this study.

The structure (Figure 3A) shows the 8MG core mimicking the hydrogen bonding network of the substrate. While electron density was strong for the 8MG moiety, it was markedly weaker for the benzyl moiety (Figure S4) and its orientation is somewhat ambiguous, although the benzyl moiety appears to interact with residues F190 and K221. The weak density for this group is indicative of mobility and an absence of specific interactions, which may help to explain the flat SAR observed for the 8MG-benzyl series. The site beyond the 8MG-binding pocket is rather open (depending on loop conformations), such that S⁸-functionalized derivatives with appropriate flexibility can generally avoid clashes with the enzyme.

S⁸-methylenecarbonyl derivatives exhibit improved binding. 8MG-acetophenone derivatives, for example **32** ($K_D = 3.9 \mu\text{M}^{34}$), were previously shown to display higher affinities for EcDHPS than 8MG (Table 4).³⁴ As an acetyl derivative also showed improved affinity ($K_D = 7.8 \mu\text{M}^{34}$), this suggested that the carbonyl group was primarily responsible for the increased potency. Various amidic compounds with alkyl and benzylic *N*-substitution were therefore tested (Table 4), a simple *N*-methyl acetamide group (**37**, $K_D = 4.5 \mu\text{M}$) proving the most beneficial. Extending off the methylene linker between the S and amide group was found to significantly reduce affinity (**43**, $K_D > 200 \mu\text{M}$).

X-ray structures of EcDHPS bound to **32** and **37** reveal a hydrogen bond between the carbonyl group of these compounds and the guanidinium group of R255 (Figure 3). The amide of **37** and phenyl group of **32** overlay well and are oriented approximately perpendicular to the plane of the guanidinium group. The phenyl ring of **32** makes a favorable van der Waals interaction with the alkyl sidechain of K221, though it should be noted that electron density is noticeably poorer for this portion

of the molecule, possibly indicating mobility. The *p*-brominated variant of **32**, **33**, bound with a K_D of 2.0 μ M.

Substituting a secondary amide for a carboxylic acid (**37** vs. **45**) gave a marked improvement in affinity ($K_D = 830$ nM). Addition of an extra methylene unit between the S and carboxylic acid group (**44**) reduced binding by 10-fold. Aryl substitution at the methylene linker of **45** (**46**) maintained affinity ($K_D = 650$ nM), contrasting with the effect of adding a methyl substituent to the methylene linker in the *N*-methyl acetamide-functionalized compound, **43**.

Crystal structures of EcDHPS in complex with **45** and **46** were solved, revealing that the carbonyl groups in the carboxylic acid moieties are positioned similarly to the carbonyl groups within the aforementioned structures, and that the benzyl group of **46** projects towards loop 2 (Figure 4). In the **46**-bound complex, the interacting loop 2 residues differ between the monomers, with the key residues being T62 for monomer A and R63/P64 for monomer B. The racemic form of **46** was synthesized and tested, however only the (*R*)-enantiomer was present in the X-ray structure, suggesting that the affinity reported in Table 4 is underestimated up to 2-fold. The (*S*)-enantiomer would likely clash with residues F190 and K221. The interactions involving the carboxylic acid groups in these two complexes are not dissimilar to those observed for the carboxylic acid-functionalized pterin analogues reported previously by Zhao et al.³¹

Pteric acid/EcDHPS complex. To further our understanding of the key interactions within the EcDHPS binding sites, we crystallized and solved (2.04 Å) the structure of EcDHPS complexed with pteric acid, the oxidized analogue of the catalysis product (Figure 5A). A single pteric acid molecule is bound in the active site for each monomer, although the quality of electron density differs (Figure S4). The binding of pteric acid is similar to that observed in the equivalent *Yersinia pestis*,¹⁰ *Bacillus anthracis*¹⁷ and *Coxiella burnetii*²³ complexes, with the pterin moiety forming hydrogen bonds to residues D96, N115, D185, G217, and K221, the pABA moiety interacting with the

sidechains of F190 (tyrosine in *C. burnetii*) and K221, and the carboxylate hydrogen bonding with the backbone and sidechain of S222 (Figure 5).

Recently, a series of pterin-sulfa conjugates were reported, with one of these crystallized in the binding site of *Yersinia pestis* DHPS.³⁹ To probe the potential for 8MG to act as a core from which to develop equivalent sulfa conjugates, we used the above pteric acid/EcDHPS complex as a guide to design potential compounds. Compound **47**, incorporating an ethylene linker between 8MG and a *p*-phenylsulfonamide moiety, was synthesized and then crystallized with EcDHPS. This compound was found to bind as predicted — the phenyl ring forms hydrophobic interactions with F190 and K221, in addition to loop 2 residues, while the sulfonamide group hydrogen bonds to the backbone and sidechain S222 (Figure 5B). Strong electron density is observed for the entire ligand. The nitrogen atom of the sulfonamide group is exposed, representing a handle from which to develop pteric acid-like conjugates.

DSF assay for ligand binding. A DSF assay was initially employed to examine compound binding to EcDHPS (Table 4). Surprisingly, the parent 8MG provided no increase in melt temperature at a concentration five times higher than the K_D value, while the hit compounds **45** and **46** gave the largest ΔT_m values of 5.8 and 7.2 °C, respectively. The other ligands tested displayed intermediate ΔT_m values of 1.0–4.4 °C.

In vitro antibacterial assay. The antibacterial activities of the compounds were tested in an M9 minimal media-based antibacterial assay, suitable for sulfonamide testing, against *E. coli* K-12 MG1655 cells.⁴⁰ No inhibition of bacterial growth was observable for any of the tested compounds (at 256 μ M), with poor permeation into cells being a likely reason for the lack of activity. The membrane-disrupting antibiotic colistin⁴¹ was thus used in conjunction with the compounds, leading to observation of a synergistic antibacterial effect. When tested alone, colistin had an MIC of 0.5 μ g·mL⁻¹. A concentration of 0.25 μ g·mL⁻¹ reduced cell growth by ~50% after 16 h and was thus chosen as a suitable level to test alongside the inhibitors (inhibitors had minimal efficacy when

colistin was used below this concentration). With colistin present, most compounds demonstrated some efficacy, however bacterial growth, although delayed, was visible after 16 h at the top concentration (256 μM) tested for many of these. The best inhibition (with colistin present) was observed for compounds **18**, **25**, **26**, **33**, **38**, **39**, and **46**, displaying MIC values of 128 μM , while **9**, **31**, and **47** displayed MIC values of 256 μM .

Binding to HPPK. 8MG and its derivatives have previously been shown to bind and inhibit the preceding enzyme, HPPK, of the folate pathway. Compound **46** was therefore tested against the HPPK enzyme from *E. coli* (EcHPPK) and *Staphylococcus aureus* (SaHPPK). For EcHPPK, binding of **46** was minimal ($K_D > 200 \mu\text{M}$), suggesting that the antibacterial activity observed for this compound was due principally to inhibition of the DHPS enzyme. For SaHPPK, **46** bound much more tightly, and independent of cofactor, with affinities of 2.7 μM and 3.5 μM in the absence and presence of ATP (1 mM), respectively. This contrasts with prior studies on HPPK, where binding of 8MG S⁸ derivatives was consistently found to be cooperative with saturating ATP present, and significantly worse in its absence.^{34, 35, 42} The cofactor-independent binding of this compound therefore opens possibilities for synthetic elaboration into the ATP-binding site of SaHPPK.

Compound **33** displays robust binding to both EcHPPK ($K_D = 0.70 \mu\text{M}$ ³⁵) and EcDHPS ($K_D = 2.0 \mu\text{M}$), and demonstrated the best antibacterial efficacy (MIC = 128 μM , equivalent to 48.7 $\mu\text{g}\cdot\text{mL}^{-1}$) of the compounds. It therefore represents a promising lead for the future development of inhibitors with dual HPPK/DHPS activity. From our studies, an obvious way to improve affinity for DHPS would be to methylate at the N⁹ position of our compounds, however N⁹-methylation has previously been demonstrated to markedly reduce binding to HPPK.³³ There is, of course, the potential that any developed inhibitors might inhibit other enzymes of the folate biosynthesis pathway, such as the validated antimicrobial target dihydrofolate reductase.

CONCLUSION

There is currently a need for novel inhibitors of DHPS to overcome sulfonamide resistant strains, and the pterin pocket of this well-validated target represents a druggable site worthy of increased attention. Here, we have developed a number of S⁸-functionalized 8MG derivatives that possess sub-micromolar affinities for EcDHPS, with the best performing compounds being up to 100-fold more potent than the parent compound. While the 8MG core may suffer intrinsic permeability issues due to its high polar surface area, the thiol group provides a convenient reaction handle for structural elaboration, allowing for the efficient generation of libraries of compounds. Future work will entail efforts to develop 8MG conjugates that form optimal interactions with regions adjacent to the pterin-binding site (the pyrophosphate and pABA-binding sites), as well as to identify compounds capable of targeting multiple enzymes of the folate biosynthesis pathway.

EXPERIMENTAL SECTION

Preparation of enzymes. SaHPPK was expressed and purified as reported in ref 43. EcHPPK was expressed and purified as reported in ref 35. EcDHPS was expressed and purified as reported in ref 34, with the additional step of removal of the hexahistidine tag for protein used in crystallographic studies. This was performed as follows: subsequent to elution from the Ni-NTA IMAC column (Qiagen), the protein sample was buffer exchanged to remove any imidazole present and 100 units of thrombin (Sigma-Aldrich) were then added. The sample was gently mixed overnight at 4 °C for 16 h. This method was noted to qualitatively (SDS-PAGE analysis) cleave all of the EcDHPS present. The sample was then reloaded onto the Ni-NTA IMAC column and collected, with subsequent purification using a Superdex 75 16/60 size exclusion as per ref 34. It was found that using a 1 mL Ni-NTA IMAC column, rather than the previous 5 mL column, improved the purity obtained such

that the previously reported³⁴ ion exchange purification step was unnecessary. Samples for crystallographic studies were dispensed as 50 μL aliquots, snap frozen, and stored at $-80\text{ }^{\circ}\text{C}$.

Differential scanning fluorimetry assay. A DSF assay was performed to screen for buffer conditions that promote the stability, and therefore crystallizability, of EcDHPS (Figure S5). Protein was present at a final concentration of $60\text{ }\mu\text{g}\cdot\text{mL}^{-1}$. Well volume was $20\text{ }\mu\text{L}$. Plates were heated from $25\text{--}95\text{ }^{\circ}\text{C}$ with a heating rate of $1.0\text{ }^{\circ}\text{C}\cdot\text{min}^{-1}$. The fluorescence intensity was measured with excitation/emission = $490/570\text{ nm}$ using a Bio-Rad CFX96 or CFX384 thermocycler. Data were acquired on the BioRad CFX Manager (version 3.1) and processed using Meltdown.⁴⁴

T_{m} values were calculated from the negative peak of the first derivative of the melt curve. A pH/buffer screen³⁷ was performed, with subsequent testing in a variant of the Hampton Research Solubility and Stability Screen, prepared by the CSIRO Collaborative Crystallisation Centre (C3).

The ligand binding assay (Figure S6) was performed as above in the determined EcDHPS buffer (50 mM trisodium citrate, 150 mM NaH_2PO_4 , and 350 mM K_2HPO_4 , pH 6.9) with $500\text{ }\mu\text{M}$ of compound (**44** present at $75\text{ }\mu\text{M}$ and **45** present at $250\text{ }\mu\text{M}$ due to limited DMSO solubility) to a final DMSO concentration of 2.5% [v/v]. ΔT_{m} values were obtained by subtracting the T_{m} of ligand-bound enzyme to that of apo-enzyme (2.5% [v/v] DMSO present as control).

Crystallization and X-ray structure determination. Crystallization experiments were performed at the C3 Centre. The sitting-drop vapor diffusion method was used at $8\text{ }^{\circ}\text{C}$ with droplets consisting of 150 nL protein solution and 150 nL reservoir solution, and using a reservoir volume of $50\text{ }\mu\text{L}$. After initial crystals were obtained, subsequent crystallization trials were prepared with seed stock prepared using seed-beads,⁴⁵ and droplets consisted of 150 nL protein solution, 120 nL reservoir solution, and 30 nL of seed stock. For the EcDHPS/**45** complex, apo-crystals were obtained from a reservoir solution containing 0.2 M MgCl_2 , 25% [w/v] PEG 3350, and 0.1 M sodium HEPES, pH 7.5. **45** was soaked into the crystals overnight prior to freezing. The concentration of protein was 10

mg·mL⁻¹. All other complexes were obtained through co-crystallization with a protein concentration of 11.1 mg·mL⁻¹; ligands were added to a concentration of 1 mM, with samples briefly centrifuged prior to dispensing to remove any precipitate. As the guanine moiety of our ligands absorbs UV light, UV imaging was used to rank crystals; crystals that appear darker under UV imaging having ostensibly better ligand occupancy (Figure S7). The reservoir solution for EcDHPS complexes were as follows: **9** – 0.147 M magnesium acetate, 27.1% [w/v] MPEG 5000, and 0.1 M tris chloride, pH 8.8; **32/46** – 0.066 M magnesium acetate, 17% [w/v] PEG 8000, and 0.1 M sodium cacodylate, pH 6.3; **37** – 0.083 M magnesium sulfate, 24.5% [w/v] PEG 6000, and 0.1 M sodium cacodylate, pH 6.3; **47** – 0.156 M magnesium chloride, 28.6% [w/v] MPEG 5000, 0.1 M tris chloride, pH 7.7; pterioic acid – 0.134 M magnesium sulfate, 19.7% [w/v] PEG 8000, and 0.1 M sodium cacodylate, pH 6.0. The protein solution contained 50 mM trisodium citrate, 150 mM NaH₂PO₄, and 350 mM K₂HPO₄ at pH 6.9.

Data were collected at the MX-1⁴⁶ or MX-2 beamline of the Australian Synchrotron using an ADSC Quantum 210 or 315 detector, respectively; 360 frames were obtained with a one degree oscillation angle for a complete data set. These data were indexed using XDS⁴⁷ and scaled using Aimless.⁴⁸ 1AJ2¹⁵ was used to solve the initial phases of the complexes by molecular replacement using Phaser.⁴⁹ Refinement was performed using REFMAC5⁵⁰ or Phenix⁵¹ and the electron density maps were visualized in Coot.⁵² After several rounds of manual rebuilding, ligands and water molecules were added and the models further refined. PDB_REDO⁵³ was used to inform refinement parameters. Data collection and refinement statistics are summarized in Table 5, and fully presented in Supplementary Table S1. Electron density maps for ligands are presented in Figure S4.

Surface plasmon resonance (SPR). SPR experiments were performed using a Biacore T200 biosensor (GE Healthcare) or ProteOn XPR36 (Bio-Rad) — Tables 1–4 state the instrument used for each compound. Experiments performed on the Biacore T200 biosensor were performed as described previously.³⁴ ProteOn XPR36 experiments were performed using a GLH chip. EcHPPK testing was

performed as per ref 35. EcDHPS testing was performed as follows: following chip preconditioning with HCl, NaOH, and SDS, EcDHPS was coupled to the surface at 298 K in HBS-P+ running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% [v/v] Tween-20) after activation with a 1:1 mixture of NHS/EDC (*N*-hydroxysuccinimide/*N*-ethyl-*N'*-(3-diethylaminopropyl)carbodiimide). Protein was diluted to $\sim 200 \mu\text{g}\cdot\text{mL}^{-1}$ at pH 5.2 and injected for 6 min. 6-methylamino-5-nitrosoisocytosine was present at 1 μM in the injection mixture to protect the pterin-binding site. The surface was then blocked with 1 M ethanolamine, pH 8.0, with an immobilization of $\sim 17,000$ RU. GRP-78 (78 kDa glucose-regulated protein) was immobilized as a control protein; protein was diluted to $58 \mu\text{g}\cdot\text{mL}^{-1}$ at pH 5.0 and injected for 5 min. The surface was then blocked as above with GRP-78 immobilization levels of $\sim 16,400$ RU. SPR experiments were performed at 20 °C in a buffer containing 50 mM HEPES, pH 8.0, 150 mM NaCl, 3% DMSO, 0.05% Tween, and 5mM DTT.

In all experiments analytes were serially diluted (2-fold or 3-fold) in SPR binding buffer and injected for 30 s contact time at $60 \mu\text{L}\cdot\text{min}^{-1}$ (Biacore) or 60 s contact time at $30 \mu\text{L}\cdot\text{min}^{-1}$ (ProteOn), then allowed to dissociate for 60 s. Separate analyte samples were prepared when performing replicates, however replicates were typically performed on the same chip. Compounds that were retested on a separate chip are noted in the Supporting Information. Compounds **32**, **35**, **37**, and **45** were independently tested on both Biacore and ProteOn instruments from discrete samples as a measure of inter-instrument variability. Binding sensorgrams were processed, solvent-corrected, and double-referenced using Scrubber software (BioLogic Software, Australia). SPR binding analysis of most compounds in this study did not display sufficiently slow dissociation rates to allow global fitting to a kinetic binding model. Typically the k_d (dissociation rate constant) must be $< 0.5 \text{ s}^{-1}$ for SPR instruments to collect sufficient amount of binding points in the dissociation phase. Therefore, to determine binding affinities (K_D values), responses at equilibrium for each analyte were fitted to a 1:1 steady-state affinity model available within Scrubber as previously described.³³ SPR sensorgrams and steady-state binding isotherms are presented in Figures S1–3.

In vitro antibacterial assay. MIC values for inhibitors were determined as per ref 40 using M9 media and *E. coli* K-12 MG1655 cells. When tested alongside the inhibitors, colistin was present at a concentration of 0.25 $\mu\text{g}\cdot\text{mL}^{-1}$.

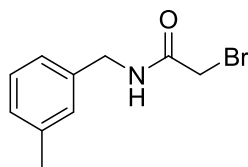
Compound Procurement and Analysis. 8MG and **32** were purchased from Sigma-Aldrich and TimTec respectively. Pterioic acid was purchased from Schircks Laboratories.

Compounds **1–8** were synthesized and characterized as described in ref 33. Compound **34** was synthesized and characterized as described in ref 34. Compounds **9–27**, **32**, **33**, **35–37**, and **40–43** were synthesized and characterized as described in ref 35. Other compounds were synthesized as described below. In all cases, ^1H NMR spectra were recorded on a 400 MHz Bruker NMR spectrometer and chemical shifts were referenced to the residual proteo-solvent peak. Chemical shifts are reported as δ values in parts per million (ppm), and coupling constants (J) in Hertz (Hz). Exchangeable protons at the N^3 , N^7 , N^9 , or O^4 positions can potentially be observed as broad signals at high frequency in the ^1H NMR spectrum (> 10 ppm) and their observation depends on the protonation state and potential intermolecular interactions of the guanine.⁵⁴ In the parent compound, 8MG, two signals at 12.43 and 10.79 ppm are observed and the latter was assigned as a NH from a SOFAST ^{15}N HMQC experiment. The broad signals > 10 ppm in the ^1H NMR spectra of the compounds below are reported if observed but have not been specifically assigned to NH or OH. Liquid-chromatography-mass spectrometry (LCMS) was acquired on an Agilent UHPLC/MS (1260/6120) system using a Poroshell 120 EC-C18 column (3.0 mm x 50 mm, 2.7 micron) with a flow rate of 0.5 $\text{mL}\cdot\text{min}^{-1}$. Elution was achieved with standard HPLC buffers (buffer A: 99.9% H_2O /0.1% formic acid; buffer B: 99.9% CH_3CN /0.1% formic acid) using a gradient from 5% B/95% A to 100% B over 5 min. Preparative reverse-phase HPLC was performed on an Agilent 1260 Prep HPLC using an Alltima C8 column (250 mm x 22 mm, 5 micron) with a flow rate of 20 $\text{mL}\cdot\text{min}^{-1}$. Elution was achieved using preparative HPLC buffers (buffer C: 99.9% H_2O /0.1% TFA; buffer D:

99.9% CH₃CN/0.1% TFA) using a gradient from 5% D/95% C to 95% D over 20 min. Final compounds were determined to be > 95% pure by analytical HPLC and ¹H NMR.

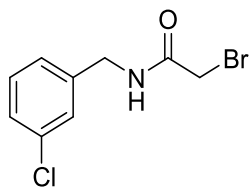
Synthesis of precursors 48 and 49. Bromoacetyl bromide (1 equ.) was added drop-wise to a stirring solution of the respective benzylamine (1 equ.) and DIPEA (2 equ.) in anhydrous DCM at 0 °C. The solution was allowed to warm to rt and stirred overnight. The reaction was diluted with additional DCM and washed with brine and water. The organic layer was dried with anhydrous MgSO₄ and concentrated under reduced pressure. The resulting residue was purified by silica flash chromatography (DCM/MeOH) to yield the products as brown solids.

2-Bromo-*N*-(3-methylbenzyl)acetamide (48)



Yield: 51%. ¹H NMR (400 MHz, CDCl₃) δ 7.24 (m, 1H, ArH), 7.10 (m, 3H, ArH), 6.76 (bs, 1H, NH), 4.44 (d, *J* = 5.8 Hz, 2H, NHCH₂), 3.92 (s, 2H, CH₂Br), 2.35 (s, 3H, CH₃). ¹³C NMR (100 MHz, CDCl₃) δ 165.4 (C=O), 138.7 (C), 137.3 (C), 128.9 (CH), 128.7 (CH), 128.7 (CH), 124.9 (CH), 44.3 (CH₂), 29.3 (CH₂), 21.5 (CH₃).

2-Bromo-*N*-(3-chlorobenzyl)acetamide (49)



Yield: 44%. ^1H NMR (400 MHz, CDCl_3) δ 7.27 (m, 3H, ArH), 7.16 (m, 1H, ArH), 6.80 (bs, 1H, NH), 4.44 (d, $J = 6.0$ Hz, 2H, CH_2NH), 3.93 (s, 2H, CH_2Br). ^{13}C NMR (100MHz, CDCl_3) δ 165.7 (C=O), 139.5 (C), 134.8 (C), 130.2 (CH), 128.1 (CH), 127.9 (CH), 125.9 (CH), 43.7 (CH_2), 29.1 (CH_2).

General procedure for the alkylation of 8MG. 8MG (0.1 g, 0.55 mmol) was dissolved in 0.5 M NaOH solution (3 mL) and to the stirring solution alkylating reagent (0.6 mmol) was added dropwise. The reaction was allowed to proceed overnight at rt. The reaction mixture was then made mildly acidic by the dropwise addition of acetic acid resulting in the precipitation of an off-white amorphous solid. The solid was filtered and dried to give each compound with yields of 27–94%. Compound **47** was further purified by preparative reverse-phase HPLC (vide supra). Fractions containing pure product were lyophilized to give **47** as a white solid.

2-Amino-8-((naphthalen-1-ylmethyl)thio)-1,9-dihydro-6H-purin-6-one (28)

Yield: 27%. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 12.60 (bs, 1H, NH), 10.61 (bs, 1H, NH), 8.18 (d, $J = 8.3$ Hz, 1H, ArH), 7.95 (d, $J = 7.4$ Hz, 1H, ArH), 7.85 (d, $J = 8.2$ Hz, 1H, ArH), 7.61–7.48 (m, 3H, ArH), 7.43–7.39 (m, 1H, ArH), 6.33 (bs, 2H, NH_2), 4.89 (s, 2H, linker CH_2). LCMS: 324.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((naphthalen-2-ylmethyl)thio)-1,9-dihydro-6H-purin-6-one (29)

Yield: 51%. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 11.08 (bs, 1H, NH), 7.83 (m, 4H, ArH), 7.52 (m, 1H, ArH), 7.47 (m, 2H, ArH), 6.52 (bs, 2H, NH_2), 4.56 (s, 2H, linker CH_2). LCMS: 324.2 $[\text{M} + \text{H}]^+$

2-Amino-8-(phenylethynyl)-1,9-dihydro-6H-purin-6-one (30)

Yield: 74%. ^1H NMR (400 MHz, DMSO) δ 12.94 (bs, 1H), 10.66 (bs, 1H), 7.59 (m, 2H, ArH), 7.48 (m, 3H, ArH), 6.47 (bs, 2H, NH_2). LCMS: 252.1 $[\text{M} + \text{H}]^+$.

2-Amino-8-((pyridin-4-ylmethyl)thio)-1,9-dihydro-6H-purin-6-one (31)

Yield: 73%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.55 (bs, 1H, NH), 10.58 (bs, 1H, NH), 8.53 (dd, *J* = 4.7, 1.4 Hz, 2H, ArH), 7.46 (d, *J* = 6.0 Hz, 2H, ArH), 6.33 (bs, 2H, NH₂), 4.42 (s, 2H, linker CH₂). LCMS: 275.1 [M + H]⁺.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N-(3-methylbenzyl)acetamide (38)

Yield: 46% ¹H NMR (400 MHz, DMSO) δ 10.81 (bs, 1H), 8.66 (t, *J* = 5.7 Hz, 1H, NH), 7.17 (t, *J* = 7.7 Hz, 1H, ArH), 7.02 (m, 3H, ArH), 6.58 (bs, 2H, NH₂), 4.25 (d, *J* = 5.8 Hz, 2H, NHCH₂Ar), 3.96 (s, 2H, SCH₂), 2.24 (s, 3H, CH₃). LCMS: 345.1 [M+H]⁺.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-N-(3-chlorobenzyl)acetamide (39)

Yield: 94% ¹H NMR (400 MHz, DMSO) δ 12.51 (bs, 1H), 10.48 (bs, 1H), 8.74 (t, *J* = 5.7 Hz, 1H, NH), 7.34–7.21 (m, 4H, ArH), 6.30 (bs, 2H, NH₂), 4.29 (d, *J* = 6.0 Hz, 2H, NHCH₂), 3.94 (s, 2H, SCH₂). LCMS: 365.0 [M+H]⁺ (³⁵Cl), 367.0 [M+H]⁺ (³⁷Cl) (3:1).

3-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)propanoic acid (44)

Yield: 78% ¹H NMR (400 MHz, DMSO) δ 12.49 (bs, 1H), 10.53 (bs, 1H), 6.25 (bs, 2H, NH₂), 3.26 (t, *J* = 6.9 Hz, 2H, CH₂), 2.66 (t, *J* = 6.9 Hz, 2H, CH₂). LCMS: 255.9 [M+H]⁺.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)acetic acid (45)

Yield: 50%. ¹H NMR (400 MHz, DMSO) δ 10.82 (bs, 1H), 6.29 (bs, 2H, NH₂), 3.34 (s, 2H, CH₂). LCMS: 242.1 [M+H]⁺.

2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)-2-phenylacetic acid (46) Yield: 51%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 13.03 (bs, 1H, COOH), 12.70 (bs, 1H, NH), 10.55 (bs, 1H, NH), 7.45–7.41 (m, 2H, ArH), 7.39–7.29 (m, 3H, ArH), 6.29 (bs, 2H, NH₂), 5.51 (s, 1H, CH) LCMS: 318.1 [M + H]⁺.

4-(2-((2-Amino-6-oxo-6,9-dihydro-1H-purin-8-yl)thio)ethyl)benzenesulfonamide (47)

Yield: 50%. ¹H NMR (400 MHz, DMSO) δ 10.84 (bs, 1H), 7.75 (d, *J* = 8.4, 2H, ArH), 7.45 (d, *J* = 8.4 Hz, 2H, ArH), 7.29 (bs, 2H, NH₂), 3.44 (t, *J* = 7.3 Hz, 2H, CH₂), 3.05 (t, *J* = 7.3 Hz, 2H, CH₂). LCMS: 367.1 [M+H]⁺.

Supporting Information. Sensorgrams and binding isotherms corresponding to the SPR performed are found in Figure S1–3. Figures of the ligand difference ($mF_o - DF_c$) density for the crystal structures reported are found in Figure S4. Melt curves for the DSF assays are found in Figures S5 and S6. UV imaging as a tool for identifying ligand binding is found in Figure S7. Crystallographic data collection and refinement statistics are found in Table S1. MIC values for control antibiotics are found in Table S2.

PDB ID Codes for EcDHPS complexes: **9** – 5U12, **32** – 5U13, **37** – 5U11, **45** – 5U0Y, **46** – 5U0Z, **47** – 5U14, pteric acid – 5U10.

CORRESPONDING AUTHOR INFORMATION

██

ACKNOWLEDGEMENT. All crystals were grown at the C3 Crystallisation Centre at CSIRO, Parkville, Australia and X-ray data were obtained at the MX1 and MX2 beamlines at the Australian Synchrotron, Victoria, Australia. We thank OpenEye Scientific Software for a license to use their software. We thank Janet Newman for help with crystallization, and Shane Seabrook for help with DSF. JBB was supported by an NHMRC Research Fellowship 1020411.

ABBREVIATIONS: DHPS, dihydropteroate synthase; EcDHPS, dihydropteroate synthase from *E. coli*; BaDHPS, dihydropteroate synthase from *B. anthracis*; 8MG, 8-mercaptopguanine; SPR, Surface Plasmon Resonance; pABA, *p*-aminobenzoic acid; DHPPP, 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate; DHPT, dihydropteroate; MPEG, methoxy poly(ethylene glycol) GRP-78, 78 kDa glucose-regulated protein; rmsd, root-mean-square deviation.

REFERENCES

1. Manyando, C.; Njunju, E. M.; D'Alessandro, U.; Van Geertruyden, J.-P. Safety and efficacy of co-trimoxazole for treatment and prevention of *Plasmodium falciparum* malaria: a systematic review. *PLoS One* **2013**, *8*, e56916.
2. Forgacs, P.; Wengenack, N. L.; Hall, L.; Zimmerman, S. K.; Silverman, M. L.; Roberts, G. D. Tuberculosis and trimethoprim-sulfamethoxazole. *Antimicrob. Agents Chemother.* **2009**, *53*, 4789–4793.
3. Martin, S. I.; Fishman, J. A. Pneumocystis pneumonia in solid organ transplantation. *Am. J. Transplant.* **2013**, *13*, 272–279.
4. Campbell, M. L.; Marchaim, D.; Pogue, J. M.; Sunkara, B.; Bheemreddy, S.; Bathina, P.; Pulluru, H.; Chugh, N.; Wilson, M. N.; Moshos, J.; Ku, K.; Hayakawa, K.; Martin, E. T.; Lephart, P. R.; Rybak, M. J.; Kaye, K. S. Treatment of methicillin-resistant *Staphylococcus aureus* infections with a minimal inhibitory concentration of 2 µg/mL to vancomycin: old (trimethoprim/sulfamethoxazole) versus new (daptomycin or linezolid) agents. *Ann. Pharmacother.* **2012**, *46*, 1587–1597.
5. Woods, D. D. The relation of p-aminobenzoic acid to the mechanism of the action of sulphanilamide. *Br. J. Exp. Pathol.* **1940**, *21*, 74–90.
6. Brown, G. M. The biosynthesis of folic acid. II. Inhibition by sulfonamides. *J. Biol. Chem.* **1962**, *237*, 536–540.
7. Roland, S.; Ferone, R.; Harvey, R. J.; Styles, V. L.; Morrison, R. W. The characteristics and significance of sulfonamides as substrates for *Escherichia coli* dihydropteroate synthase. *J. Biol. Chem.* **1979**, *254*, 10337–10345.
8. Swedberg, G.; Castensson, S.; Sköld, O. Characterization of mutationally altered dihydropteroate synthase and its ability to form a sulfonamide-containing dihydrofolate analog. *J. Bacteriol.* **1979**, *137*, 129–136.
9. Patel, O.; Satchell, J.; Baell, J.; Fernley, R.; Coloe, P.; Macreadie, I. Inhibition studies of sulfonamide-containing folate analogs in yeast. *Microb. Drug Resist.* **2003**, *9*, 139–146.

10. Yun, M. K.; Wu, Y.; Li, Z.; Zhao, Y.; Waddell, M. B.; Ferreira, A. M.; Lee, R. E.; Bashford, D.; White, S. W. Catalysis and sulfa drug resistance in dihydropteroate synthase. *Science* **2012**, *335*, 1110–1114.
11. Haruki, H.; Pedersen, M. G.; Gorska, K. I.; Pojer, F.; Johnsson, K. Tetrahydrobiopterin biosynthesis as an off-target of sulfa drugs. *Science* **2013**, *340*, 987–991.
12. Mondal, S.; Mandal, S. M.; Mondal, T. K.; Sinha, C. Structural characterization of new Schiff bases of sulfamethoxazole and sulfathiazole, their antibacterial activity and docking computation with DHPS protein structure. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2015**, *150*, 268–279.
13. Boufas, W.; Dupont, N.; Berredjem, M.; Berrezag, K.; Bechecker, I.; Berredjem, H.; Aouf, N.-E. Synthesis and antibacterial activity of sulfonamides. SAR and DFT studies. *Journal of Molecular Structure* **2014**, *1074*, 180–185.
14. Sköld, O. Sulfonamide resistance: mechanisms and trends. *Drug Resist. Updat.* **2000**, *3*, 155–160.
15. Achari, A.; Somers, D. O.; Champness, J. N.; Bryant, P. K.; Rosemond, J.; Stammers, D. K. Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. *Nat. Struct. Biol.* **1997**, *4*, 490–497.
16. Hampele, I. C.; D'Arcy, A.; Dale, G. E.; Kostrewa, D.; Nielsen, J.; Oefner, C.; Page, M. G.; Schönfeld, H. J.; Stüber, D.; Then, R. L. Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. *J. Mol. Biol.* **1997**, *268*, 21–30.
17. Babaoglu, K.; Qi, J.; Lee, R. E.; White, S. W. Crystal structure of 7,8-dihydropteroate synthase from *Bacillus anthracis*: mechanism and novel inhibitor design. *Structure* **2004**, *12*, 1705–1717.
18. Pemble, C. W. t.; Mehta, P. K.; Mehra, S.; Li, Z.; Nourse, A.; Lee, R. E.; White, S. W. Crystal structure of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase*dihydropteroate synthase bifunctional enzyme from *Francisella tularensis*. *PLoS One* **2010**, *5*, e14165.

19. Baca, A. M.; Sirawaraporn, R.; Turley, S.; Sirawaraporn, W.; Hol, W. G. J. Crystal structure of Mycobacterium tuberculosis 6-hydroxymethyl-7,8-dihydropteroate synthase in complex with pterin monophosphate: new insight into the enzymatic mechanism and sulfa-drug action. *J. Mol. Biol.* **2000**, *302*, 1193–1212.
20. Bagautdinov, B.; Kunishima, N. Crystal Structure of dihydropteroate synthase (FolP) from Thermus thermophilus HB8. ; 2006. *RIKEN Structural Genomics/Proteomics Initiative*, doi:10.2210/pdb2dqw/pdb.
21. Levy, C.; Minnis, D.; Derrick, J. P. Dihydropteroate synthase from Streptococcus pneumoniae: structure, ligand recognition and mechanism of sulfonamide resistance. *Biochem. J.* **2008**, *412*, 379–388.
22. Morgan, R. E.; Batot, G. O.; Dement, J. M.; Rao, V. A.; Eadsforth, T. C.; Hunter, W. N. Crystal structures of Burkholderia cenocepacia dihydropteroate synthase in the apo-form and complexed with the product 7,8-dihydropteroate. *BMC Struct. Biol.* **2011**, *11*, 1–9.
23. Franklin, M. C.; Cheung, J.; Rudolph, M. J.; Burshteyn, F.; Cassidy, M.; Gary, E.; Hillerich, B.; Yao, Z. K.; Carlier, P. R.; Totrov, M.; Love, J. D. Structural genomics for drug design against the pathogen Coxiella burnetii. *Proteins: Struct., Funct., Bioinf.* **2015**, *83*, 2124–2136.
24. Lawrence, M. C.; Iliades, P.; Fernley, R. T.; Berglez, J.; Pilling, P. A.; Macreadie, I. G. The three-dimensional structure of the bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/dihydropteroate synthase of Saccharomyces cerevisiae. *J. Mol. Biol.* **2005**, *348*, 655–670.
25. Hammoudeh, D. I.; Zhao, Y.; White, S. W.; Lee, R. E. Replacing sulfa drugs with novel DHPS inhibitors. *Future Med. Chem.* **2013**, *5*, 1331–40.
26. Morrison, R. W.; Mallory, W. R.; Styles, V. L. Pyrimido[4,5-c]pyridazines. 1. Cyclizations with .alpha.-keto esters. *J. Org. Chem.* **1978**, *43*, 4844–4849.
27. Morrison, R., Jr; Mallory, W.; Styles, V. EP 0000383 (A1) Pyrimido[4,5-c]pyridazines, their use in pharmaceutical compositions, process and intermediates for their preparation. 1979.

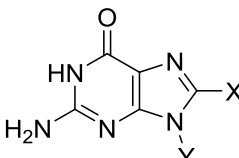
28. Lever, O. J.; Bell, L.; McGuire, H.; Ferone, R. Monocyclic pteridine analogues. Inhibition of *Escherichia coli* dihydropteroate synthase by 6-amino-5-nitrosoisocytosines. *J. Med. Chem.* **1985**, *28*, 1870–1874.
29. Lever, O. W. J.; Bell, L. N.; Hyman, C.; McGuire, H. M.; Ferone, R. Inhibitors of dihydropteroate synthase: substituent effects in the side-chain aromatic ring of 6-[[3-(aryloxy)propyl]amino]-5-nitrosoisocytosines and synthesis and inhibitory potency of bridged 5-nitrosoisocytosine-p-aminobenzoic acid analogues. *J. Med. Chem.* **1986**, *29*, 665–670.
30. Hevener, K. E.; Yun, M. K.; Qi, J.; Kerr, I. D.; Babaoglu, K.; Hurdle, J. G.; Balakrishna, K.; White, S. W.; Lee, R. E. Structural studies of pterin-based inhibitors of dihydropteroate synthase. *J. Med. Chem.* **2010**, *53*, 166–177.
31. Zhao, Y.; Hammoudeh, D.; Yun, M. K.; Qi, J.; White, S. W.; Lee, R. E. Structure-based design of novel pyrimido[4,5-c]pyridazine derivatives as dihydropteroate synthase inhibitors with increased affinity. *ChemMedChem* **2012**, *7*, 861–870.
32. Chhabra, S.; Dolezal, O.; Collins, B. M.; Newman, J.; Simpson, J. S.; Macreadie, I. G.; Fernley, R.; Peat, T. S.; Swarbrick, J. D. Structure of *S. aureus* HPPK and the discovery of a new substrate site inhibitor. *PLoS One* **2012**, *7*, e29444.
33. Chhabra, S.; Barlow, N.; Dolezal, O.; Hattarki, M. K.; Newman, J.; Peat, T. S.; Graham, B.; Swarbrick, J. D. Exploring the chemical space around 8-mercaptopguanine as a route to new inhibitors of the folate biosynthesis enzyme HPPK. *PLoS One* **2013**, *8*, e59535.
34. Dennis, M. L.; Chhabra, S.; Wang, Z.-C.; Debono, A.; Dolezal, O.; Newman, J.; Pitcher, N. P.; Rahmani, R.; Cleary, B.; Barlow, N.; Hattarki, M.; Graham, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structure-based design and development of functionalized mercaptopguanine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*. *J. Med. Chem.* **2014**, *57*, 9612–9626.
35. Dennis, M. L.; Pitcher, N. P.; Lee, M. D.; DeBono, A. J.; Wang, Z.-C.; Harjani, J. R.; Rahmani, R.; Cleary, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structural basis for the selective

- binding of inhibitors to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus* and *Escherichia coli*. *J. Med. Chem.* **2016**, *59*, 5248–5263.
36. Bourne, C. R. Utility of the biosynthetic folate pathway for targets in antimicrobial discovery. *Antibiotics* **2014**, *3*, 1–28.
37. Seabrook, S. A.; Newman, J. High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Comb. Sci.* **2013**, *15*, 387–392.
38. Dallas, W. S.; Gowen, J. E.; Ray, P. H.; Cox, M. J.; Dev, I. K. Cloning, sequencing, and enhanced expression of the dihydropteroate synthase gene of *Escherichia coli* MC4100. *J. Bacteriol.* **1992**, *174*, 5961–5970.
39. Zhao, Y.; Shadrick, W. R.; Wallace, M. J.; Wu, Y.; Griffith, E. C.; Qi, J.; Yun, M.-K.; White, S. W.; Lee, R. E. Pterin–sulfa conjugates as dihydropteroate synthase inhibitors and antibacterial agents. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 3950–4.
40. Zlitni, S.; Ferruccio, L. F.; Brown, E. D. Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. *Nat. Chem. Biol.* **2013**, *9*, 796–804.
41. Li, J.; Nation, R. L.; Milne, R. W.; Turnidge, J. D.; Coulthard, K. Evaluation of colistin as an agent against multi-resistant Gram-negative bacteria. *Int. J. Antimicrob. Agents* **2005**, *25*, 11–25.
42. Yun, M. K.; Hoagland, D.; Kumar, G.; Waddell, M. B.; Rock, C. O.; Lee, R. E.; White, S. W. The identification, analysis and structure-based development of novel inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Bioorg. Med. Chem.* **2014**, *22*, 2157–2165.
43. Chhabra, S.; Newman, J.; Peat, T. S.; Fernley, R. T.; Caine, J.; Simpson, J. S.; Swarbrick, J. D. Crystallization and preliminary X-ray analysis of 6-hydroxymethyl-7,8-dihydropterin from *Staphylococcus aureus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2010**, *66*, 575–578.
44. Rosa, N.; Ristic, M.; Seabrook, S. A.; Lovell, D.; Lucent, D.; Newman, J. Meltdown: a tool to help in the interpretation of thermal melt curves acquired by differential scanning fluorimetry. *J. Biomol. Screen.* **2015**, *20*, 898–905.

45. Luft, J. R.; DeTitta, G. T. A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55*, 988–993.
46. Cowieson, N. P.; Aragao, D.; Clift, M.; Ericsson, D. J.; Gee, C.; Harrop, S. J.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunncliffe, A.; Williamson, R.; Caradoc-Davies, T. MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **2015**, *22*, 187–190.
47. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132.
48. Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69*, 1204–1214.
49. Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 432–438.
50. Vagin, A.; Steiner, R.; Lebedev, A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2184–2195.
51. Adams, P.; Afonine, P.; Bunkóczi, G.; Chen, V.; Davis, I.; Echols, N.; Headd, J.; Hung, L.; Kapral, G.; Grosse-Kunstleve, R.; McCoy, A.; Moriarty, N.; Oeffner, R.; Read, R.; Richardson, D.; Richardson, J.; Terwilliger, T.; Zwart, P. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221.
52. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
53. Joosten, R. P.; Womack, T.; Vriend, G.; Bricogne, G. Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *65*, 176–185.

54. West, R. T.; Garza, L. A. n.; Winchester, W. R.; Walmsley, J. A. Conformation, hydrogen bonding and aggregate formation of guanosine 5'-monophosphate and guanosine in dimethylsulfoxide. *Nucleic Acids Res.* **1994**, *22*, 5128–5134.

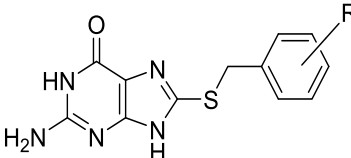
Table 1. 8MG derivatives



compd	X	Y	K_D EcDHPS (μM) ^b
1	SH	Ethyl	420 ^a
2	SH	Benzyl	350 ^a
3	SH	CH ₂ Benzyl	160 ^a
4	Br	H	42 ^a
5	Methyl	H	71 ^a
6	OH	H	200 ^a
7	OH	Me	101 ± 2
8	N-morpholine	H	160 ^a

Binding constants are presented as mean ± standard deviation. ^an = 1. ^bTested using a Biacore T200.

Table 2. Binding of 8MG-benzyl derivatives as determined by SPR

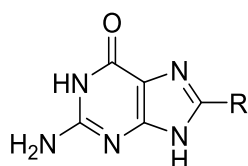


compd	R	K_D EcDHPS (μM) ^b
9	2-fluoro	8.5 ± 2.1 ^a
10	2-chloro	7.0 ± 0.8
11	2-bromo	14.0 ± 0.2
12	2-methyl	9.5 ± 0.6
13	2-cyano	12.7 ± 0.3
14	2-nitro	33 ± 2
15	3-methyl	11 ± 2
16	3-methoxy	7.5 ± 0.5

17	3-cyano	11 ± 2
18	4-fluoro	11 ± 2
19	2,2-difluoro	14 ± 3
20	2,4-difluoro	12 ± 3
21	3,4-difluoro	11 ± 1
22	2-fluoro,3-methyl	31 ± 8
23	2-methyl,4-fluoro	13 ± 1
24	3-methyl,4-fluoro	8.2 ± 0.9
25	2,3-dimethyl	14 ± 2
26	2,5-dimethyl	11.5 ± 0.0
27	3,5-dimethyl	14 ± 3

Binding constants are presented as mean ± standard deviation. ^aDennis et al.³⁴ ^bTested using a Biacore T200.

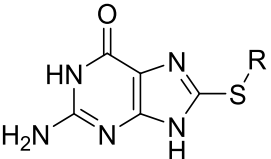
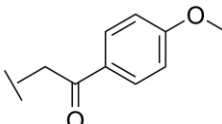
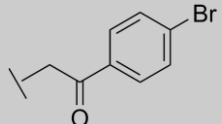
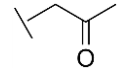
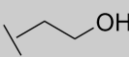
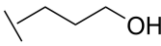
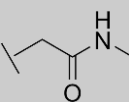
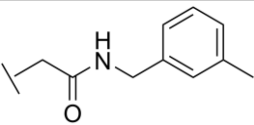
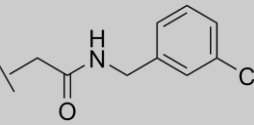
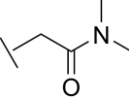
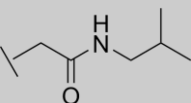
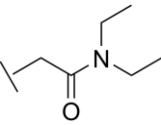
Table 3. Binding of 8MG-benzyl analogues as determined by SPR

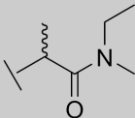
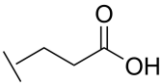
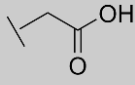
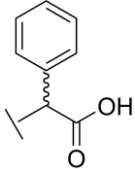
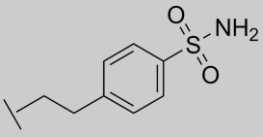


compd	R	K_D EcDHPS (μ M) ^a	K_D EcHPPK (μ M) ^{ab}
28		15 ± 3	7.3 ± 0.6
29		16 ± 2	1.46 ± 0.04
30		9.8 ± 0.9	nd
31		15 ± 1	15.9 ± 0.1

Binding constants are presented as mean ± standard deviation. ^aTested using a Biacore T200 ^bin the presence of 10 mM MgCl₂ and 1 mM ATP. nd - not determined.

Table 4. 8MG derivatives

			
compd	R	K_D EcDHPS (μM)	ΔT_m ($^\circ$) ^d
8MG	H	76.2 ± 12^{ab}	0.0 ± 0.2
32		3.9 ± 0.2^{ab} (4.3 ± 0.2^c)	1.6 ± 0.1
33		2.0 ± 0.1^c	4.9 ± 0.3
34		7.8 ± 1.0^{ab}	nd
35		18 ± 4^b (11.8 ± 0.4^c)	3.6 ± 0.4
36		20.0 ± 0.5^c	nd
37		4.5 ± 0.1^b (7.6 ± 0.3^c)	3.7 ± 0.1
38		12.1 ± 0.1^c	2.2 ± 0.2
39		9.9 ± 0.6^c	2.3 ± 0.3
40		8.2 ± 0.1^c	4.4 ± 0.1
41		7.5 ± 0.2^c	4.0 ± 0.1
42		6.9 ± 0.1^c	3.4 ± 0.1

43		>200 ^c	1.0 ± 0.1
44		7.9 ± 0.3 ^c	1.2 ± 0.3 ^e
45		0.83 ± 0.01 ^b (0.86 ± 0.01 ^c)	5.8 ± 0.4 ^f
46		0.65 ± 0.03 ^b	7.2 ± 0.2
47		4.6 ± 0.1 ^c	nd

Binding constants are presented as mean ± standard deviation. Melt temperatures are presented as mean ± estimate of variation in replicates, determined using Meltdown.⁴⁴ ^aDennis et al.³⁴ Tested using a ^bBiacore T200 or ^cProteOn XPR36. ^dDSF assay performed at ligand concentration of 500 μM except where noted. ^eTested at 75 μM. ^fTested at 250 μM. nd not determined.

Table 5. Summary of crystallographic data for EcDHPS complexes

Compound	Space group	Resolution (Å)	R _{work} /R _{free}	PDB accession code
9	C2	1.84	18.0/20.4	5U12
32	C2	1.95	18.7/21.7	5U13
37	C2	1.99	19.8/23.0	5U11
45	C2	1.88	20.6/23.8	5U0Y
46	C2	2.29	19.3/24.2	5U0Z
47	C2	1.94	20.7/25.7	5U14
Pteric acid	C2	2.04	21.0/25.1	5U10

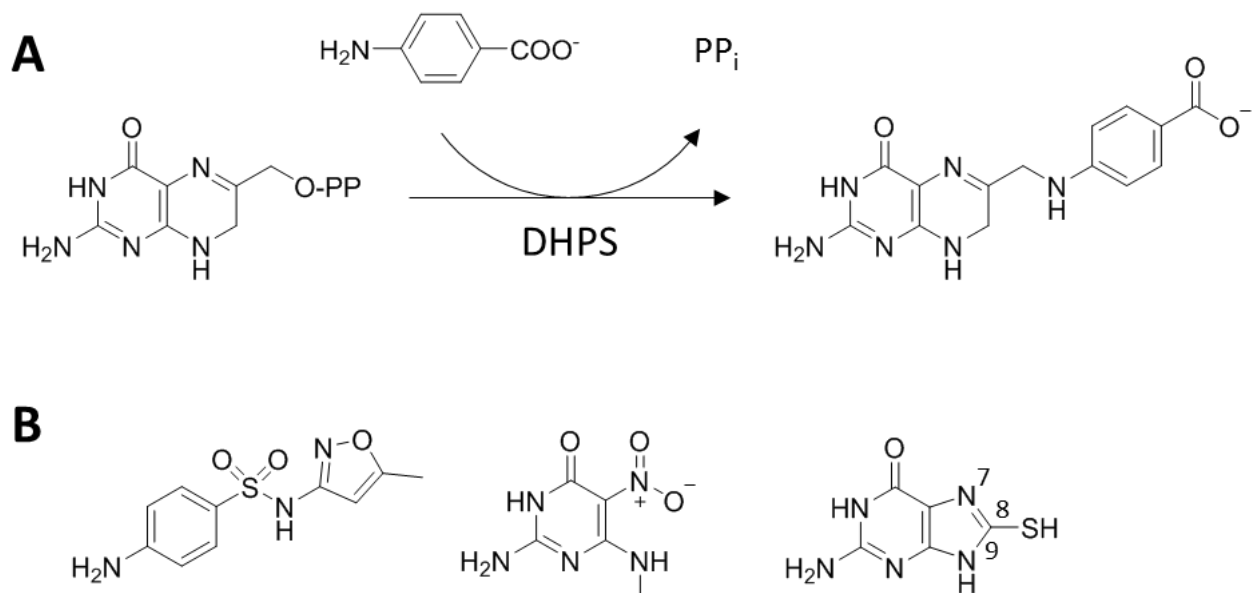


Figure 1. (A) Reaction pathway of DHPS showing the formation of DHPt from DHPPP and pABA. (B) Chemical structures of DHPS inhibitors sulfamethoxazole (left), MANIC (middle), and 8MG (right). Positions 7–9 of 8MG are labelled.

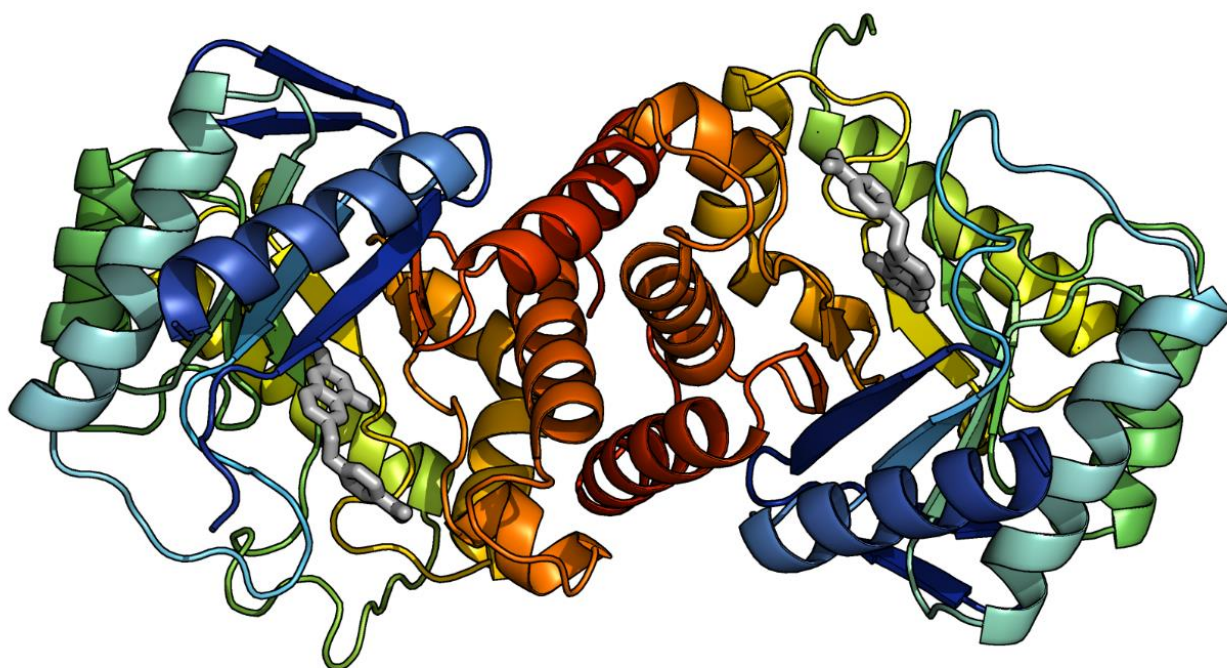


Figure 2. Structure of the EcDHPS/pteroic acid complex displayed using chainbow coloring. The pterioic acid molecules are displayed in gray.

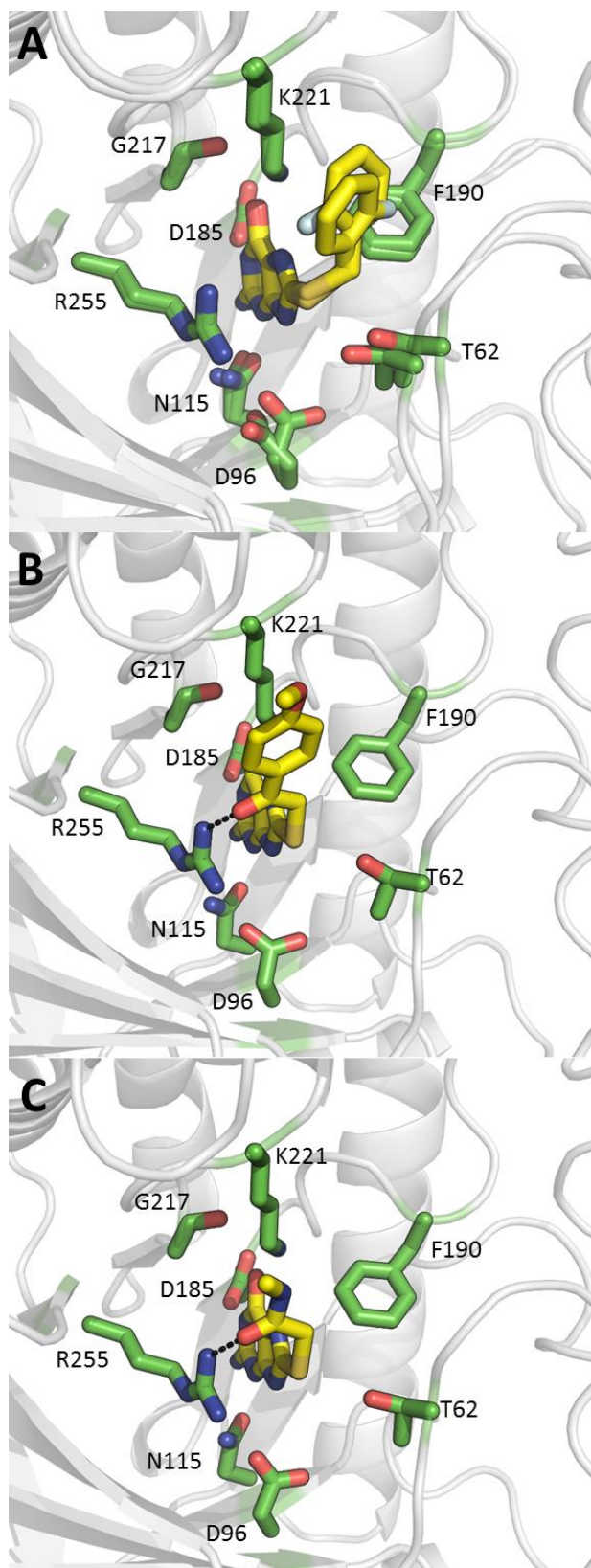


Figure 3. X-ray crystal structures of EcDHPS bound to (A) **9**, (B) **32**, and (C) **37**. Key residues are displayed in green. Hydrogen bonds to residue R255 are displayed as dashed black lines. The EcDHPS/**9** complex is displayed as a superposition of both monomers.

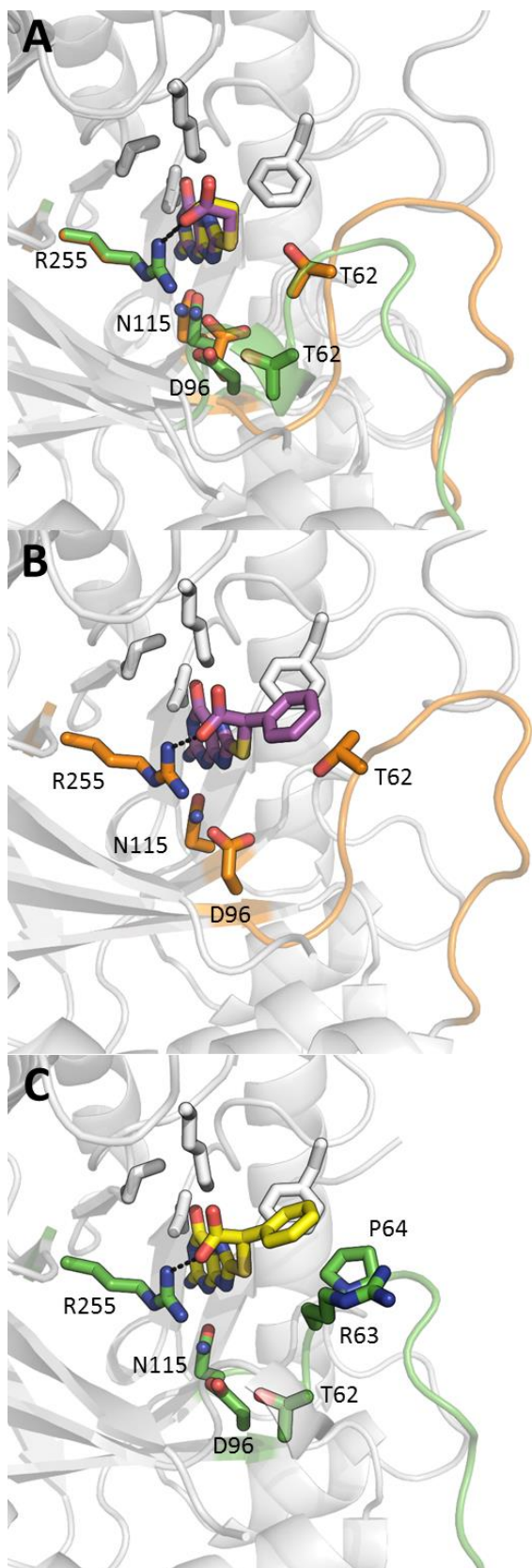


Figure 4. (A) Overlay of both monomers from the X-ray structure of **45**/EcDHPS. X-ray structure of the **46**/DHPS complex, monomer (B) A and (C) B. Key residues are displayed in orange and green, and the ligand is displayed in purple and yellow, for monomer A and B, respectively. Hydrogen bonds to residue R255 are displayed as dashed black lines.

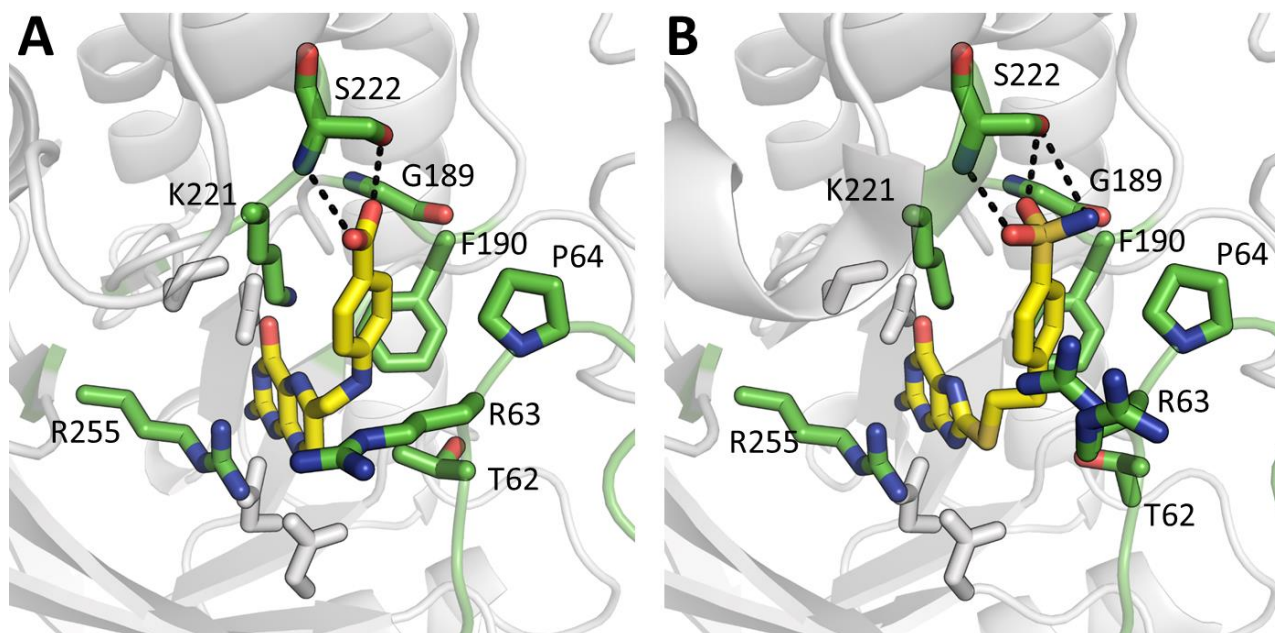


Figure 5. Crystal structure of EcDHPS bound to (A) pterioic acid or (B) **47**. Both conformations of R63 in the **47**/EcDHPS complex are displayed. Hydrogen bonds to S222 are displayed as dashed black lines.

Supporting Information

8-Mercaptoguanine derivatives as inhibitors of dihydropteroate synthase from *Escherichia coli*

*Matthew L. Dennis,^{#†} Aaron J. DeBono,[#] Michael D. Lee,[#] Noel P. Pitcher,[#] Zhong-Chang Wang,^{#‡}
Sandeep Chhabra,^{#†} Raphaël Rahmani,[#] Olan Dolezal,[†] Meghan Hattarki,[†] Bim Graham,[#] Thomas
S. Peat,[†] Jonathan B. Baell,[#] and James D. Swarbrick.[#]*

[#]Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia.

[†]CSIRO Biomedical Program, Manufacturing, Parkville, VIC 3052, Australia.

[‡]State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093,
People's Republic of China

Table of Contents

Table S1. X-ray crystallography data collection and refinement statistics for EcDHPS complexes.

Table S2. Minimum inhibitory concentration values for control antibiotics

Figure S1. Surface plasmon resonance measurements of EcDHPS.

Figure S2. Surface plasmon resonance measurements of EcHPPK

Figure S3. Surface plasmon resonance measurements of SaHPPK

Figure S4. Electron density maps of compounds within their respective EcDHPS complexes

Figure S5. Differential Scanning Fluorimetry assay of buffer conditions

Figure S6. Differential Scanning Fluorimetry assay of ligand binding

Figure S7. UV-imaging of EcDHPS crystals as an indicator of ligand binding

Table S1. Data collection and refinement statistics for the EcDHPS complexes

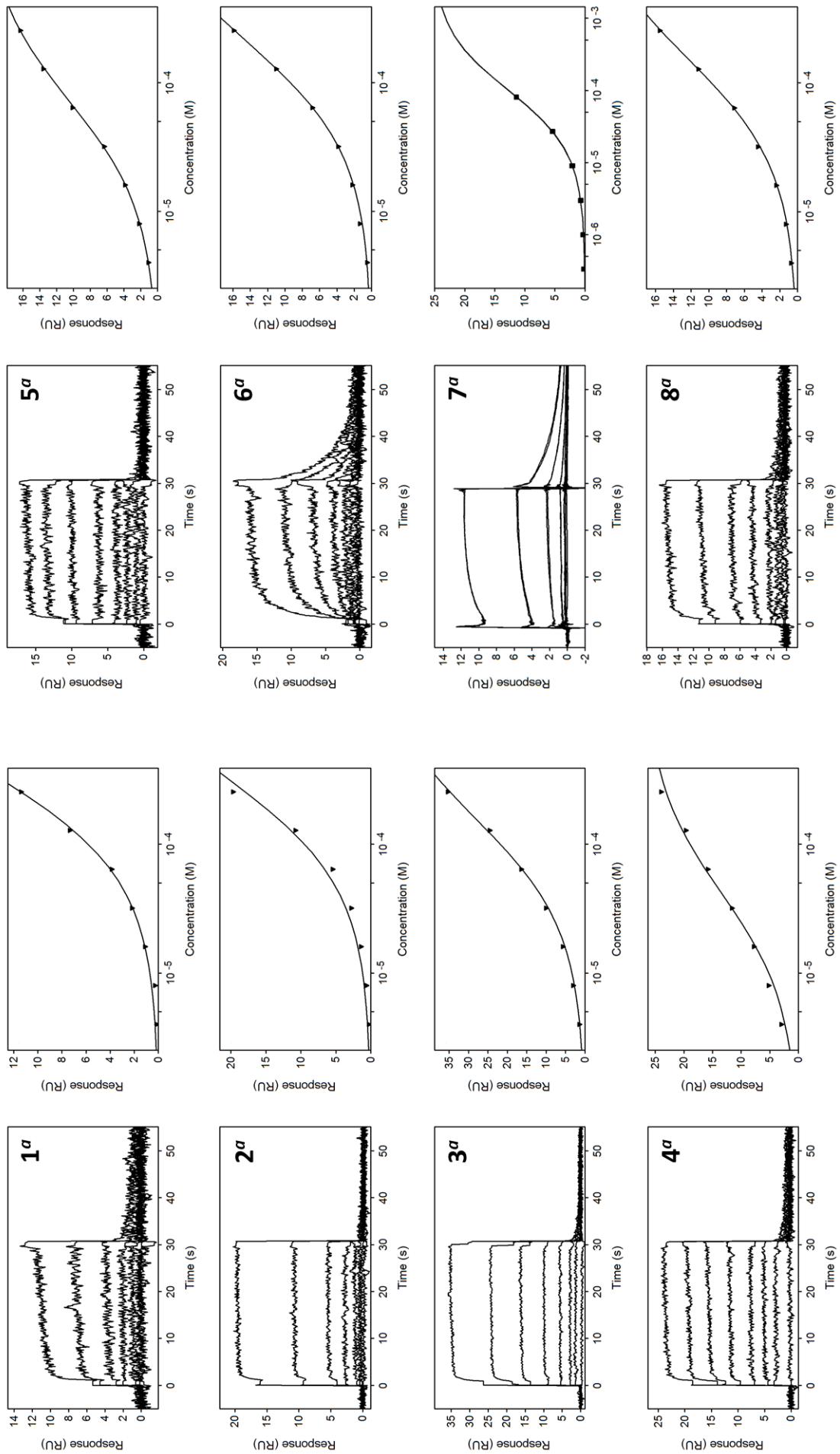
Ligand	5U12 (9)	5U13 (32)	5U11 (37)	5U0Y (45)
Space group	C2	C2	C2	C2
Wavelength (Å)	0.9537	0.9537	0.9537	0.9537
Unit-cell parameters (Å, °)	a = 95.9, b = 84.6, 84.4, α = 90, β = 109.9, γ = 90	a = 95.9, b = 84.8, c = 84.2, α = 90, β = 110.5, γ = 90	a = 94.8, b = 84.6, c = 84.0, α = 90, β = 110.5, γ = 90	a = 94.9, b = 85.2, c = 84.8, α = 90, β = 111.5, γ = 90
Diffraction data				
Resolution range (Å)	46.57-1.84 (1.88-1.84)	46.68-1.95 (2.00-1.95)	46.20-1.99 (2.04-1.99)	43.41-1.88 (1.92-1.88)
No. of unique reflections	54,783 (3,263)	45,749 (3,191)	42,177 (2,433)	51,039 (3,191)
No. of observed reflections	403,775	338,464	310,275	380,123
Completeness (%)	99.7 (97.0)	99.5 (98.3)	98.6 (81.7)	99.7 (96.6)
Data redundancy	7.4 (7.2)	7.4 (7.4)	7.4 (7.1)	7.4 (7.1)
Mean I/ σ (I)	18.6 (2.8)	18.3 (2.8)	14.8 (2.0)	27.8 (4.4)
R _{merge}	0.064 (0.702)	0.073 (0.690)	0.074 (0.791)	0.053 (0.458)
R _{p.i.m.}	0.025 (0.278)	0.029 (0.270)	0.029 (0.314)	0.021 (0.183)
Refinement				
R _{free} (%)	20.4	21.7	23.0	23.8
R _{cryst} (%)	18.0	18.7	19.8	20.6
Size of R _{free} set (%)	5.1	5.0	5.1	5.0
Protein molecules in the asymmetric unit	2	2	2	2
Inhibitor molecules	2	2	2	2
Water molecules	245	205	156	340
RMSD from ideal values				
Bond lengths (Å)	0.010	0.012	0.011	0.014
Bond angles (°)	1.28	1.40	1.34	1.50
Mean B factors (Å ²)	29.0	32.0	33.0	30.0
Ramachandran plot				
Favored (%)	98.9	99.2	98.3	99.2
Outliers (%)	0.0	0.0	0.0	0.0

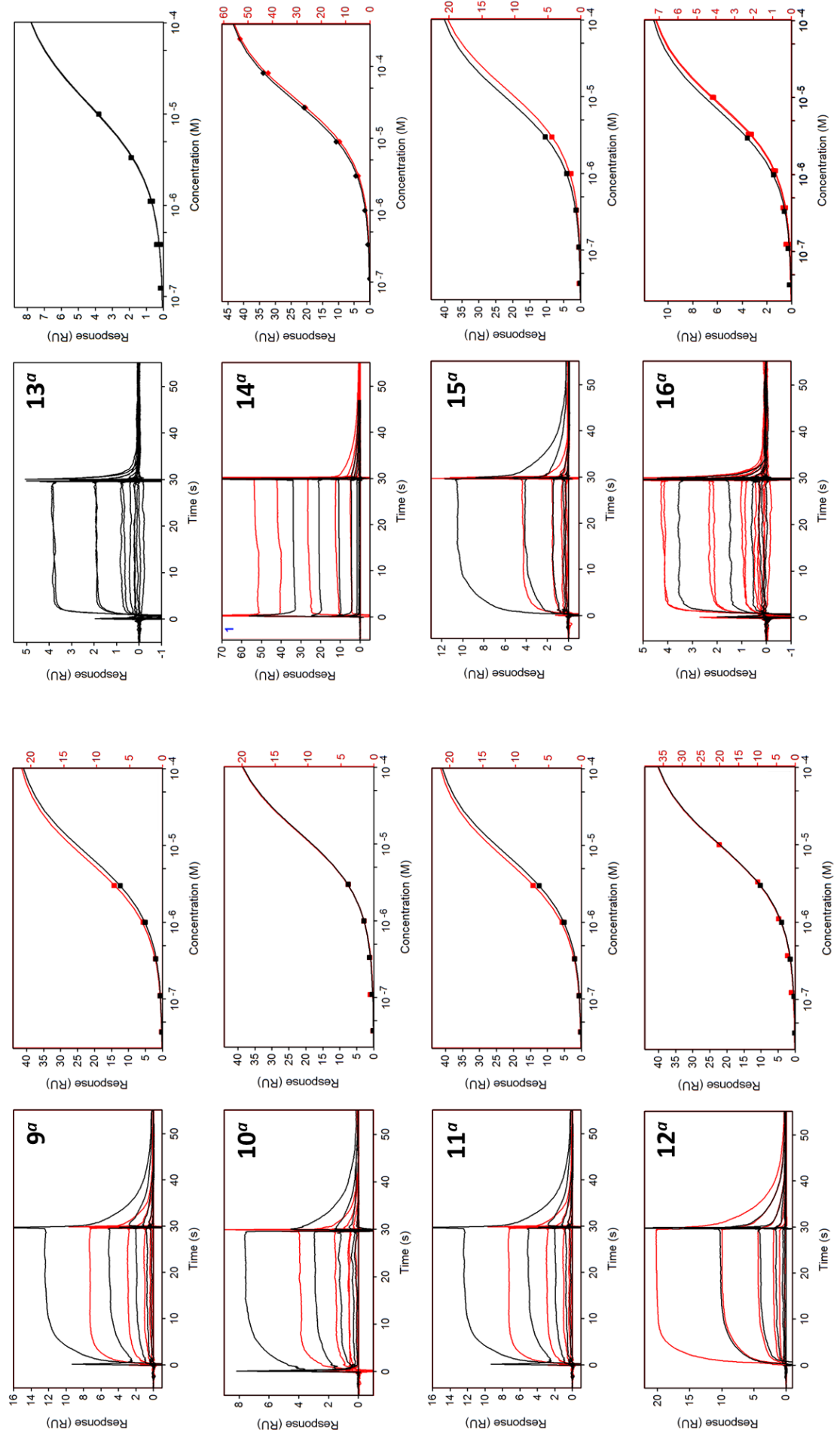
Table S1 cont.

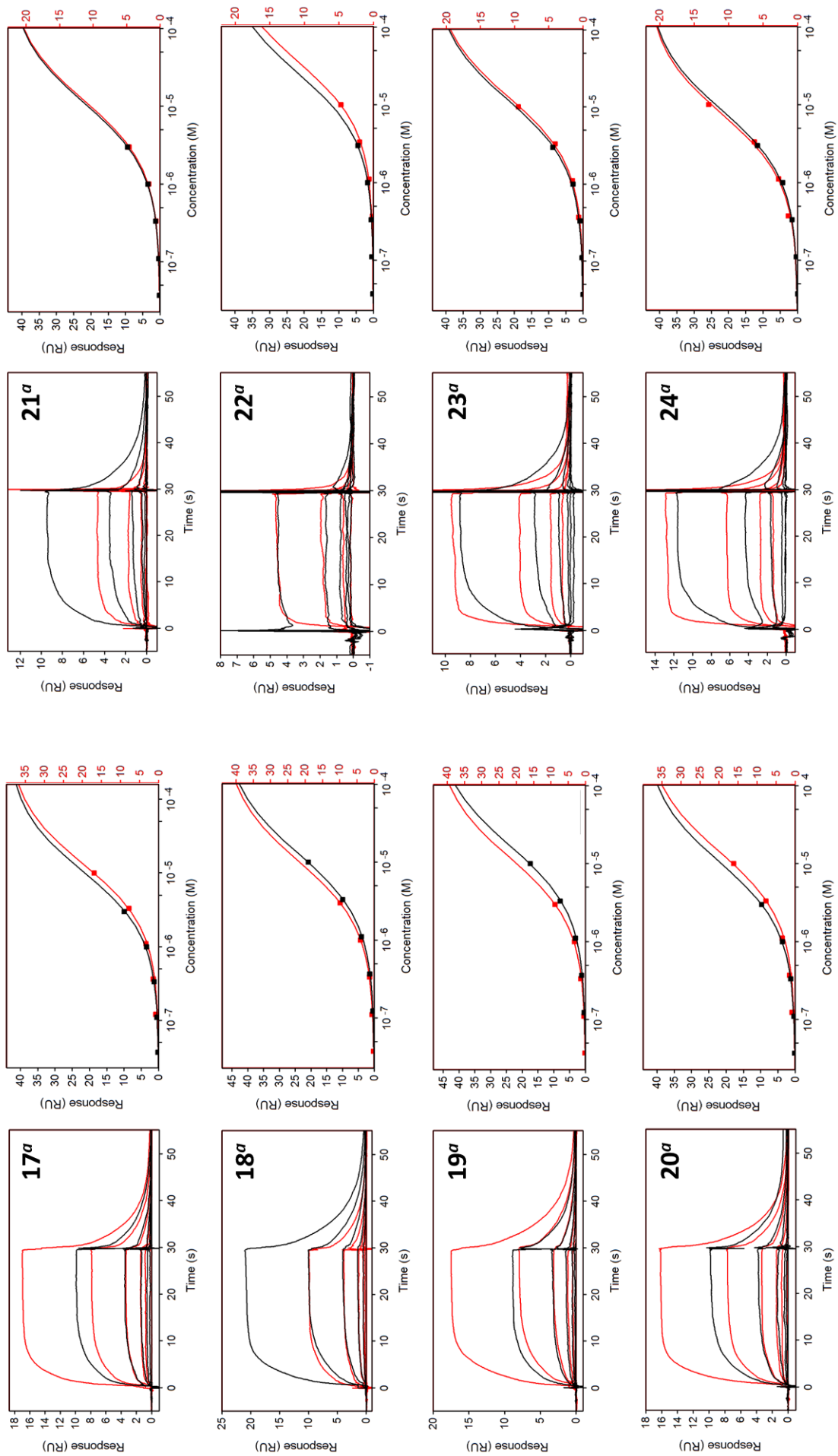
Ligand	5U0Z (46)	5U14 (47)	5U10 (Pteric acid)
Space group	C2	C2	C2
Wavelength (Å)	0.9537	0.9537	0.9537
Unit-cell parameters (Å, °)	a = 94.1, b = 85.3, c = 84.2, α = 90, β = 110.6, γ = 90	a = 94.8, b = 85.1, c = 84.1, α = 90, β = 111.1, γ = 90	a = 96.5, b = 83.9, c = 83.9, α = 90, β = 111.0, γ = 90
Diffraction data			
Resolution range (Å)	45.93-2.29 (2.37-2.29)	46.33-1.94 (1.99-1.94)	47.03-2.04 (2.09-2.04)
No. of unique reflections	27,605 (2,152)	45,137 (2,136)	39,630 (2,837)
No. of observed reflections	192,281	335,441	291,904
Completeness (%)	97.7 (77.8)	97.9 (69.2)	98.9 (86.4)
Data redundancy	7.0 (5.7)	7.4 (7.1)	7.4 (6.8)
Mean I/ σ (I)	8.6 (1.7)	29.3 (2.6)	25.4 (3.2)
R _{merge}	0.157 (0.785)	0.039 (0.683)	0.063 (0.757)
R _{p.i.m.}	0.064 (0.332)	0.023 (0.413)	0.025 (0.312)
Refinement			
R _{free} (%)	24.2	25.7	25.1
R _{cryst} (%)	19.3	20.7	21.0
Size of R _{free} set (%)	5.0	5.0	5.1
Protein molecules in the asymmetric unit	2	2	2
Inhibitor molecules	2	2	2
Water molecules	153	165	139
RMSD from ideal values			
Bond lengths (Å)	0.011	0.011	0.009
Bond angles (°)	1.39	1.30	1.33
Mean B factors (Å ²)	36.0	40.0	38.0
Ramachandran plot			
Favored (%)	98.5	98.1	98.5
Outliers (%)	0.0	0.0	0.0

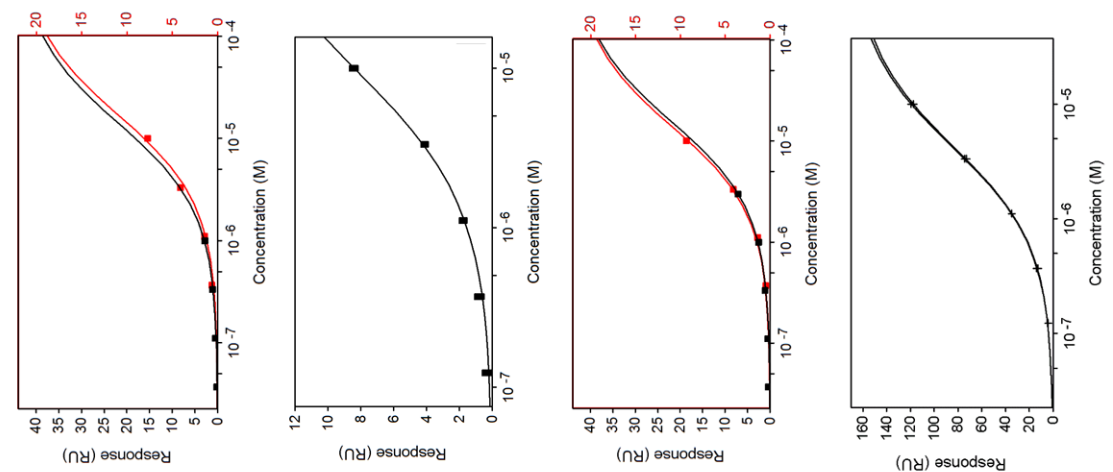
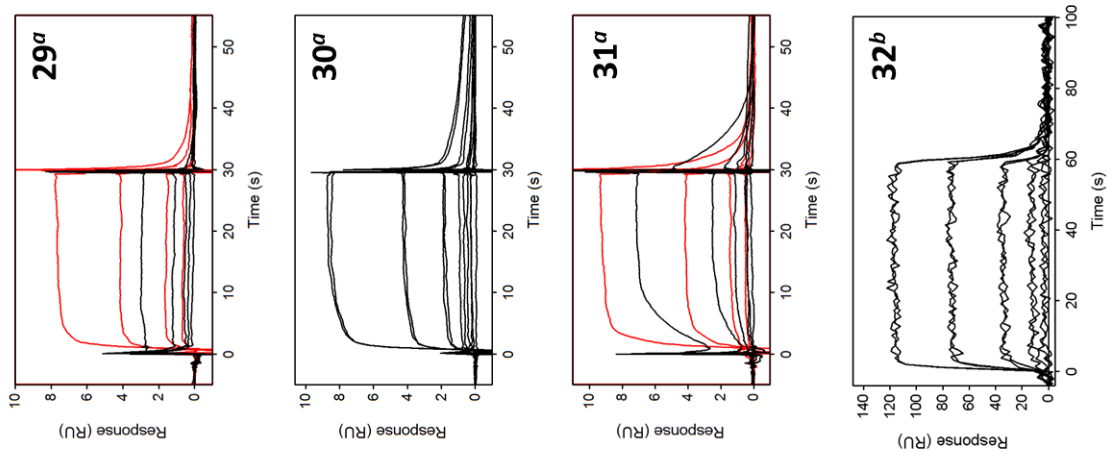
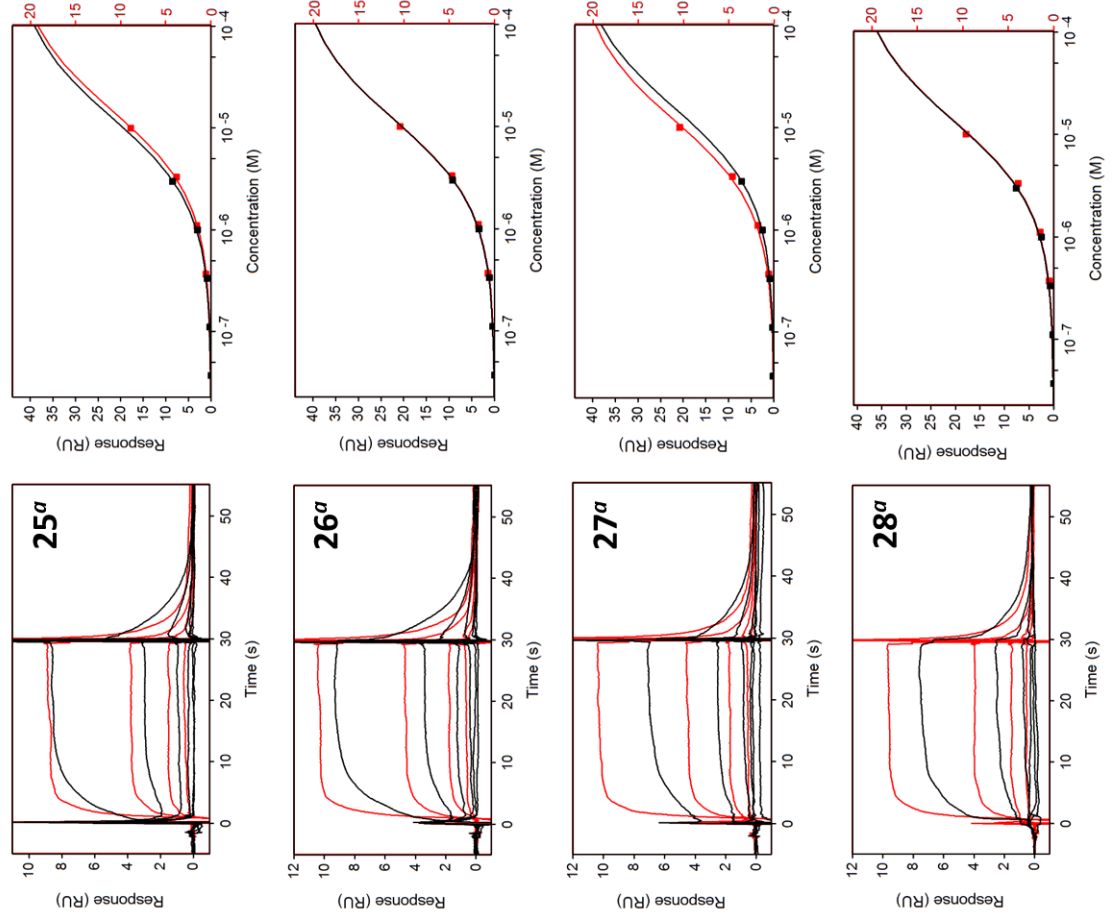
Table S2. Minimum inhibitory concentration values determined for control antibiotics

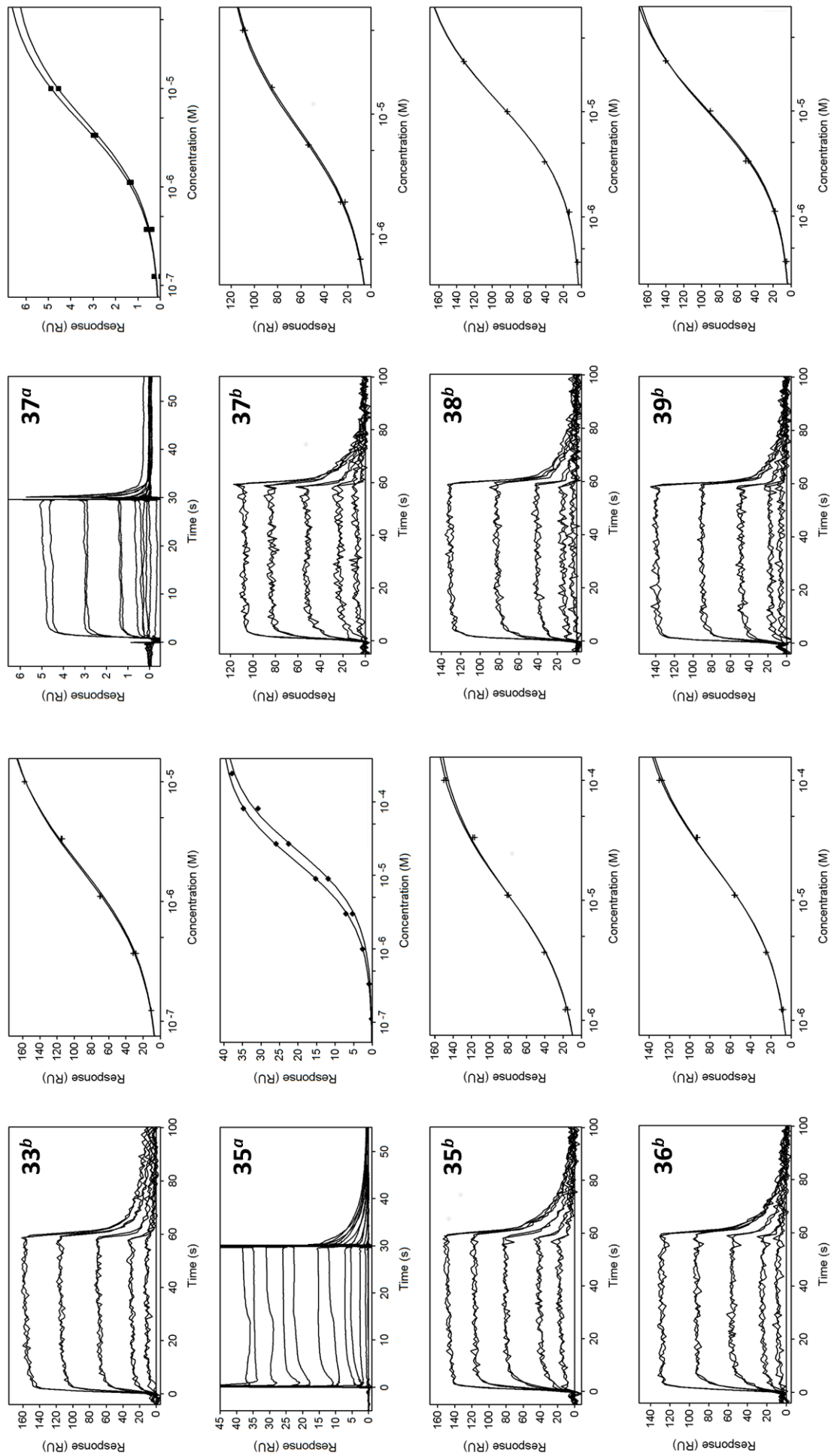
Antibiotic	MIC <i>E. coli</i> MG1655 ($\mu\text{g}\cdot\text{mL}^{-1}$)
Ampicillin	1
Trimethoprim	0.5
Polymyxin B	0.5
Colistin	0.5

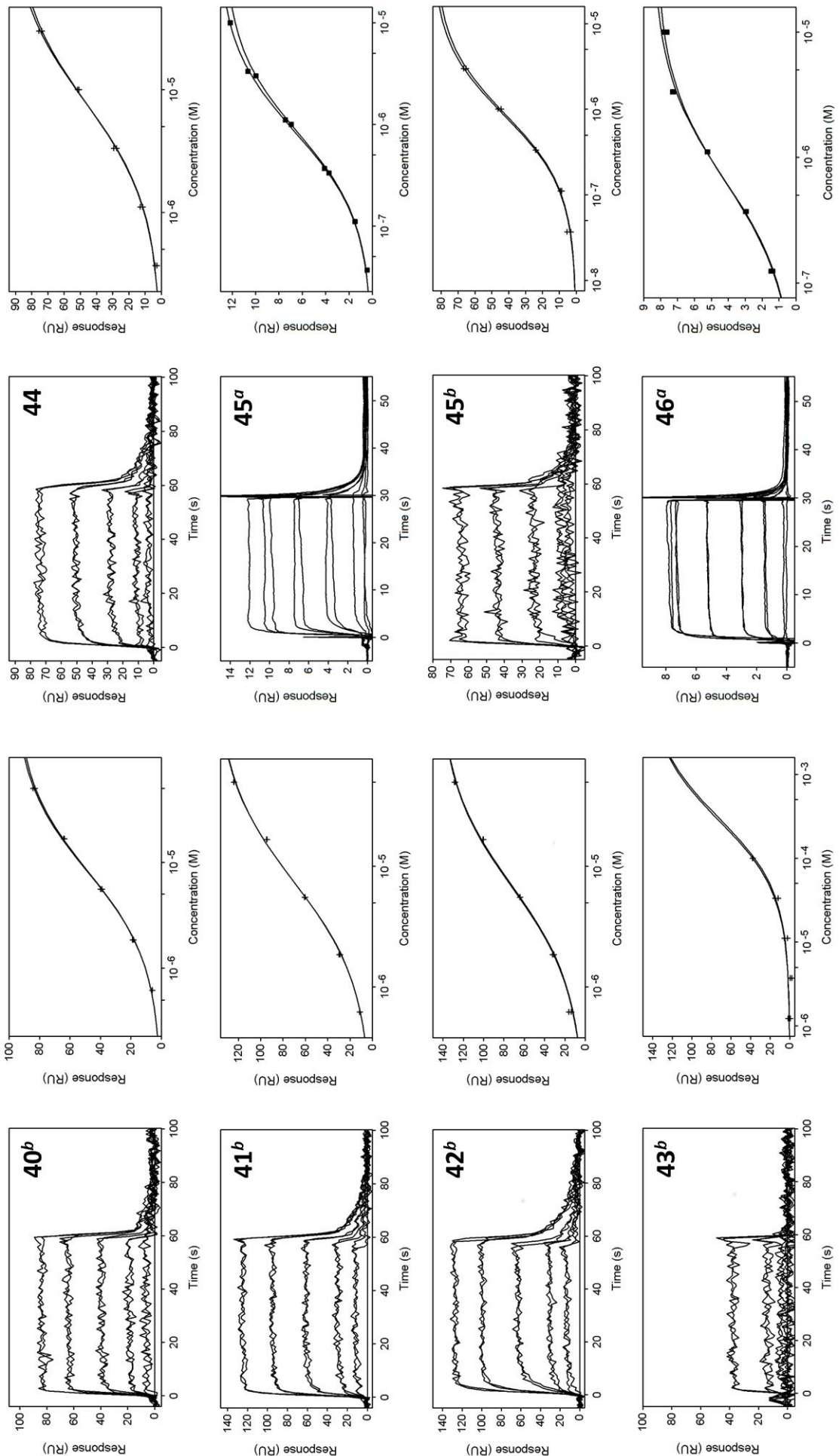












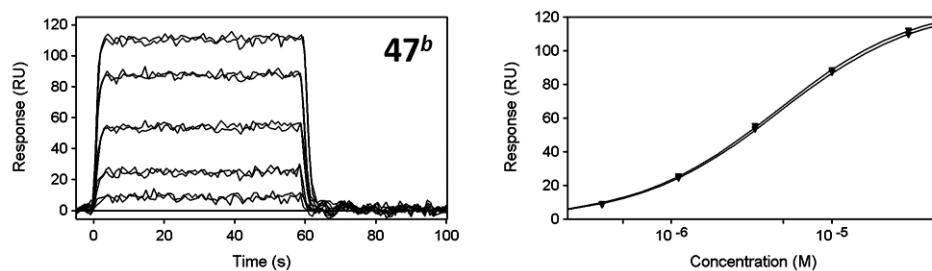


Figure S1. Surface plasmon resonance measurements of compounds **1–8**, **10–33**, and **35–47** with EcDHPS. The sensorgram (left) is displayed alongside the binding isotherms (right). For some compounds the replicates were performed on different chips (with different immobilization levels), with R_{\max} values differing greatly. To aid readers' assessment of reproducibility, the y-axis of these binding isotherms is scaled to the R_{\max} , with each replicate displayed in black or red. Compounds were tested on a ^aBiacore T200 and/or ^bProteOn XPR36 biosensor and are labelled as such.

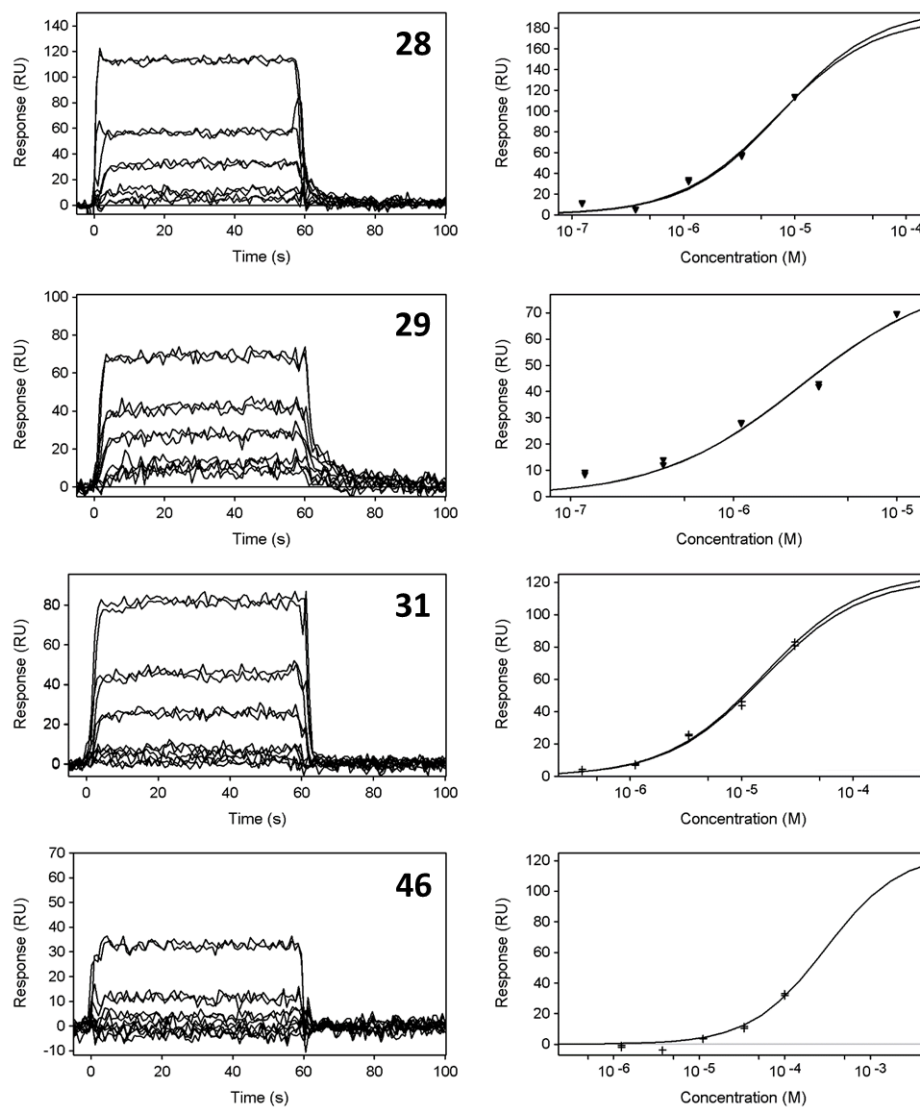


Figure S2. Surface plasmon resonance measurements of compounds **28**, **29**, **31**, and **46** with EcHPPK. Sensorgram (left) is displayed alongside the binding isotherm (right). Testing was performed on a ProteOn XPR36 biosensor (Bio-Rad).

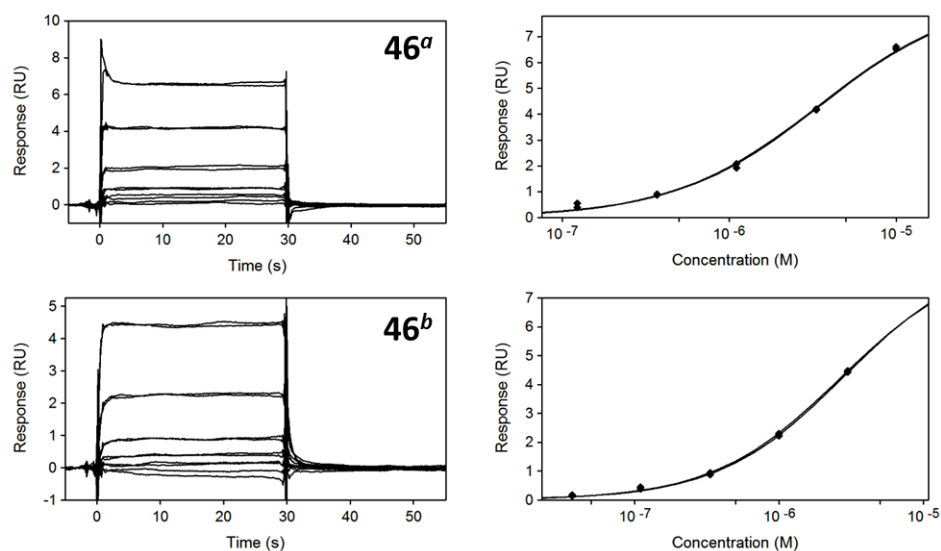
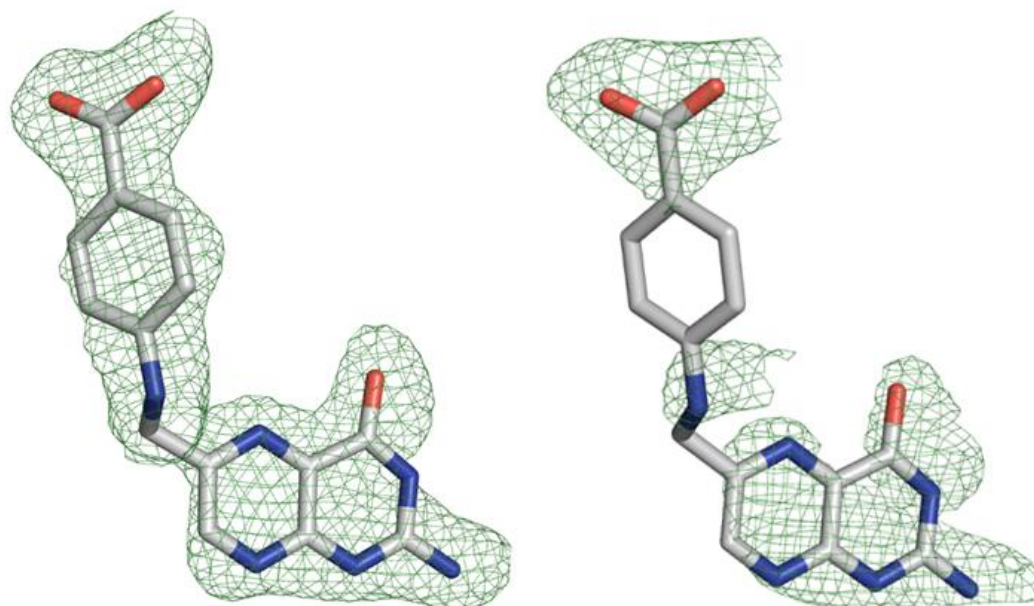
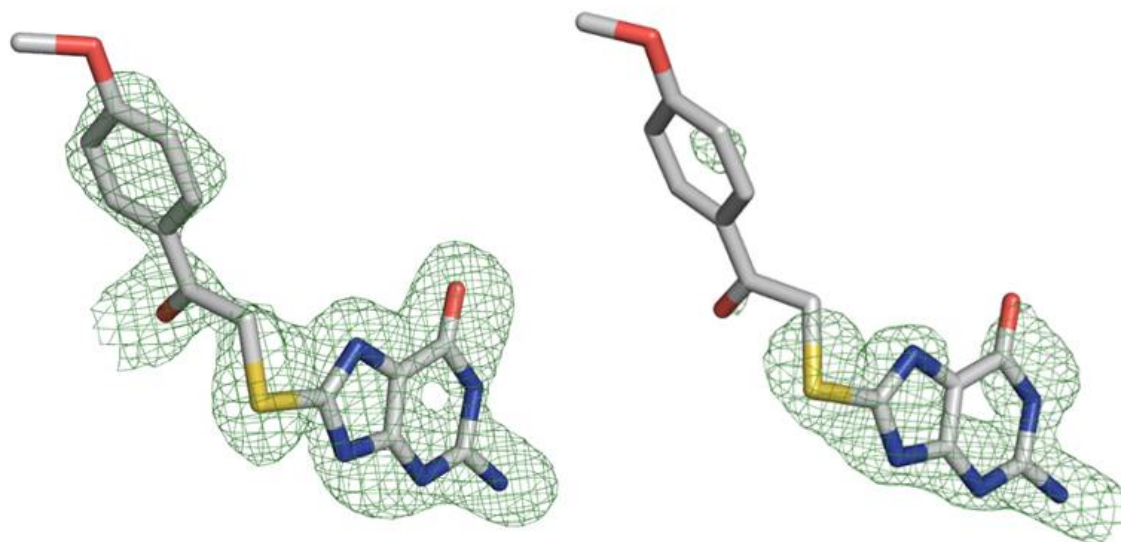


Figure S3. Surface plasmon resonance measurements of compound **46** with SaHPPK. Sensorgram (left) is displayed alongside the binding isotherm (right). Tested in the presence of ^a10 mM MgCl₂ or ^b10 mM MgCl₂ and 1 mM ATP. Testing was performed on a Biacore T200 biosensor (GE Healthcare).

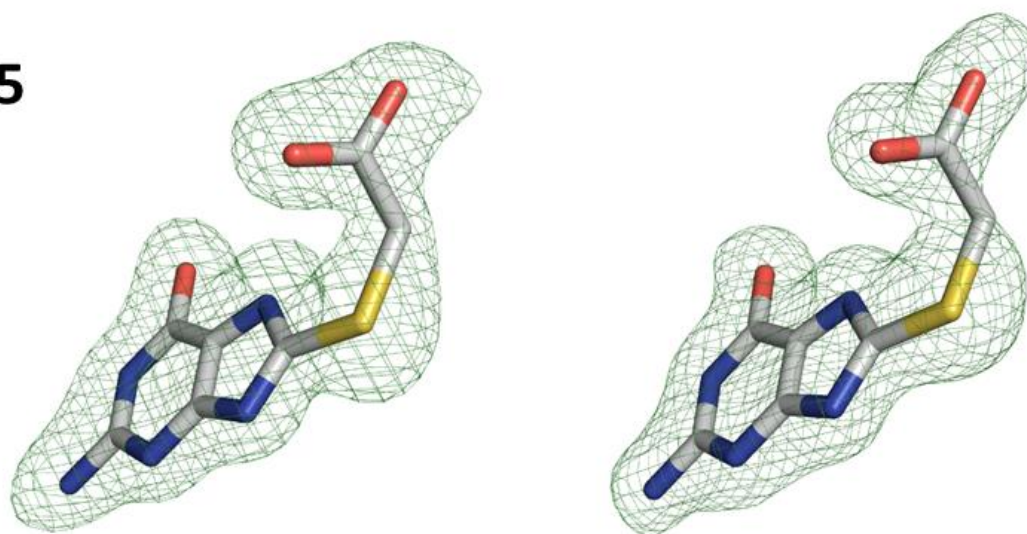
Pterioic acid



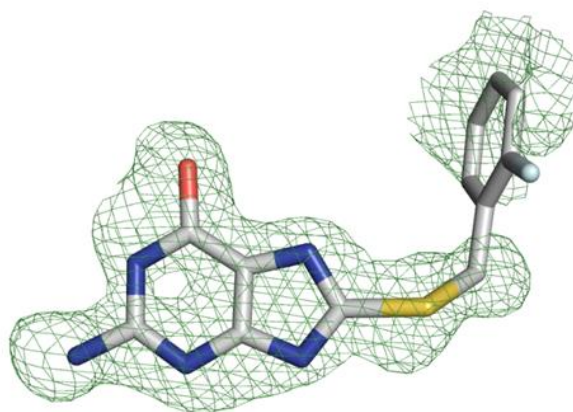
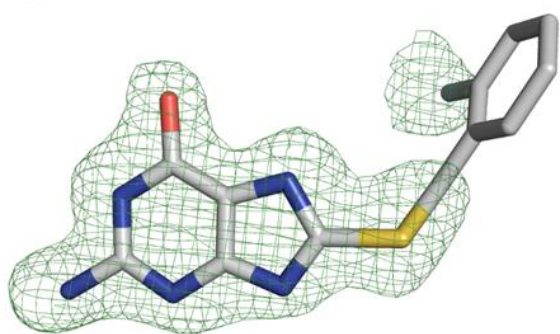
32



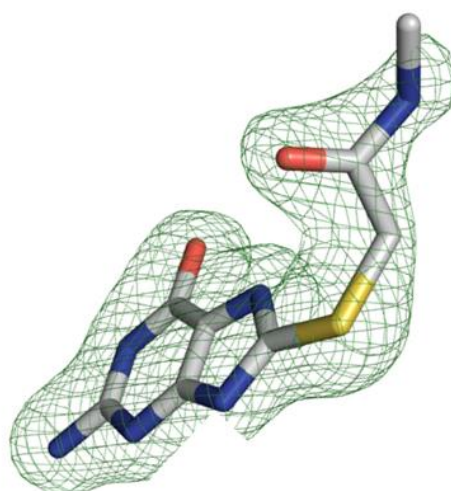
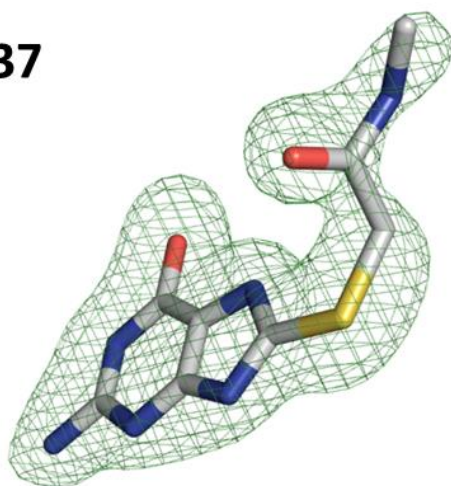
45



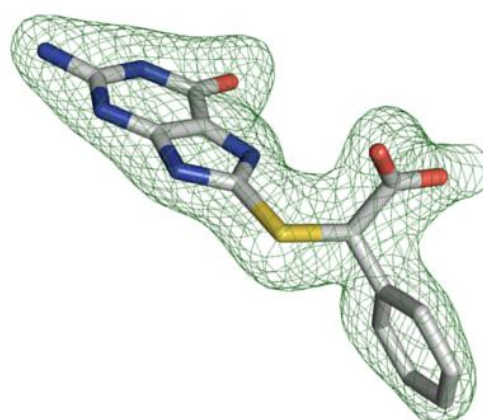
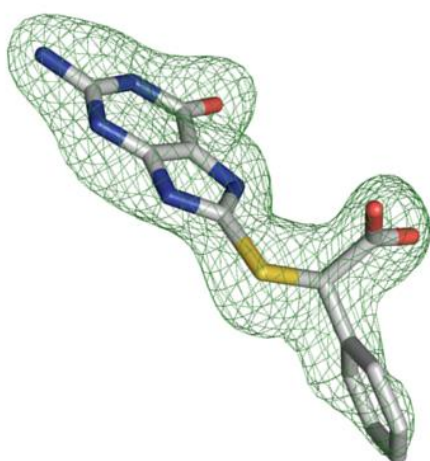
9



37



46



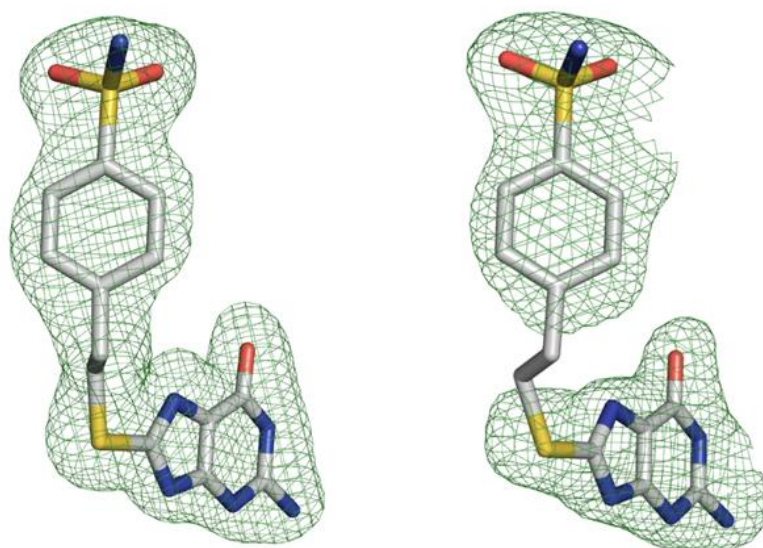


Figure S4. Electron densities of compounds in EcDHPS complexes for monomer A (left) and B (right). mF_o-DF_c difference density maps are contoured at 2.5 σ .

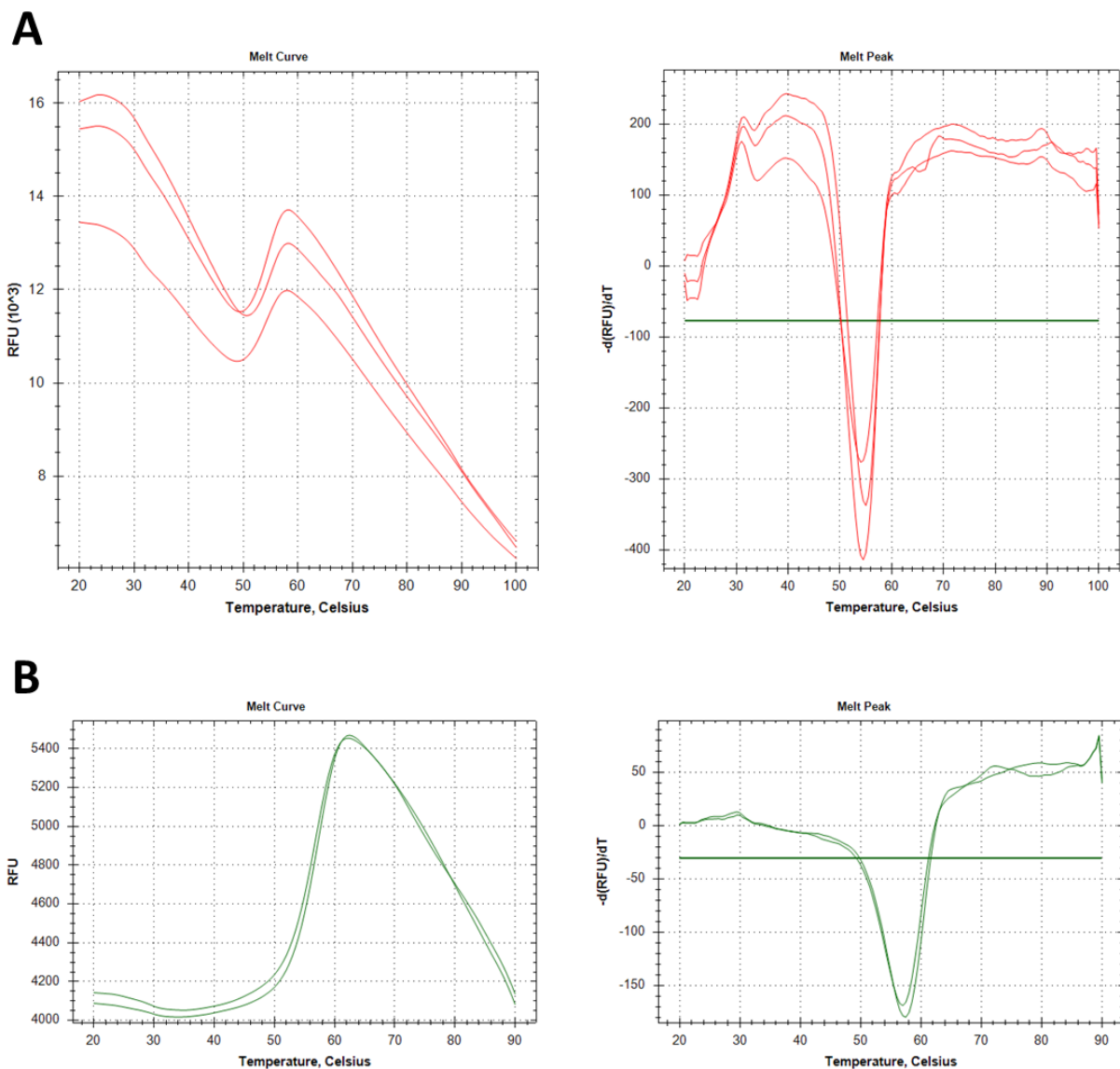


Figure S5. Melt curve (left) and first derivative of the melt curve (right) for a DSF assay of EcDHPS in a buffer containing (A) 50 mM Tris-HCl, 2 mM MgCl_2 , pH 8.5 (B) 50 mM $\text{Na}_3\text{Citrate}$, 150 mM NaH_2PO_4 , 350 mM K_2HPO_4 , pH 6.9.

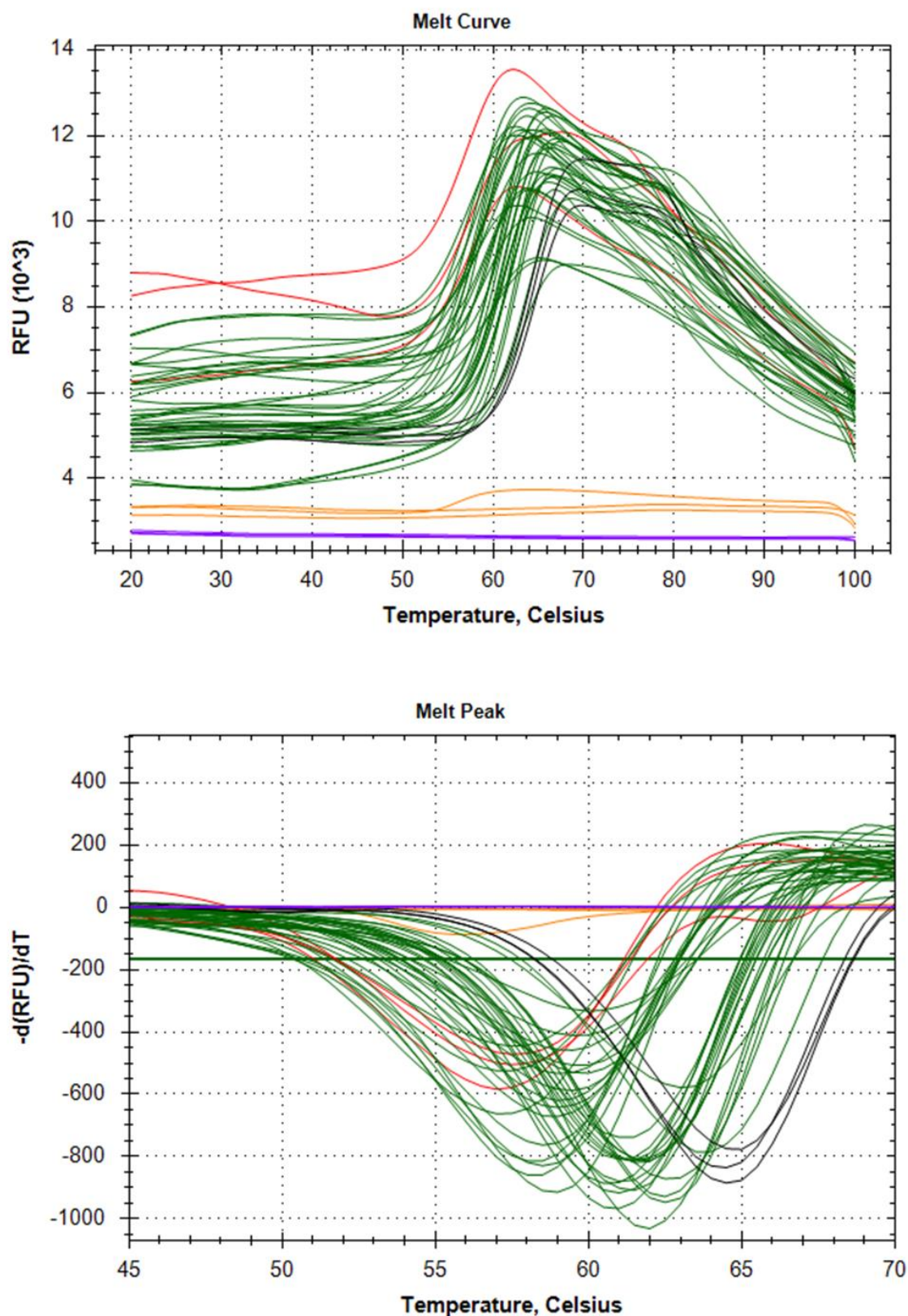


Figure S6. Melt curve (top) and first derivative of the melt curve (bottom) for a DSF assay of EcDHPS with various ligands (performed in triplicate). Controls lacking either dye or protein are displayed in purple and orange, respectively. The apo-enzyme and **46**-bound enzyme are displayed in red and black, respectively.

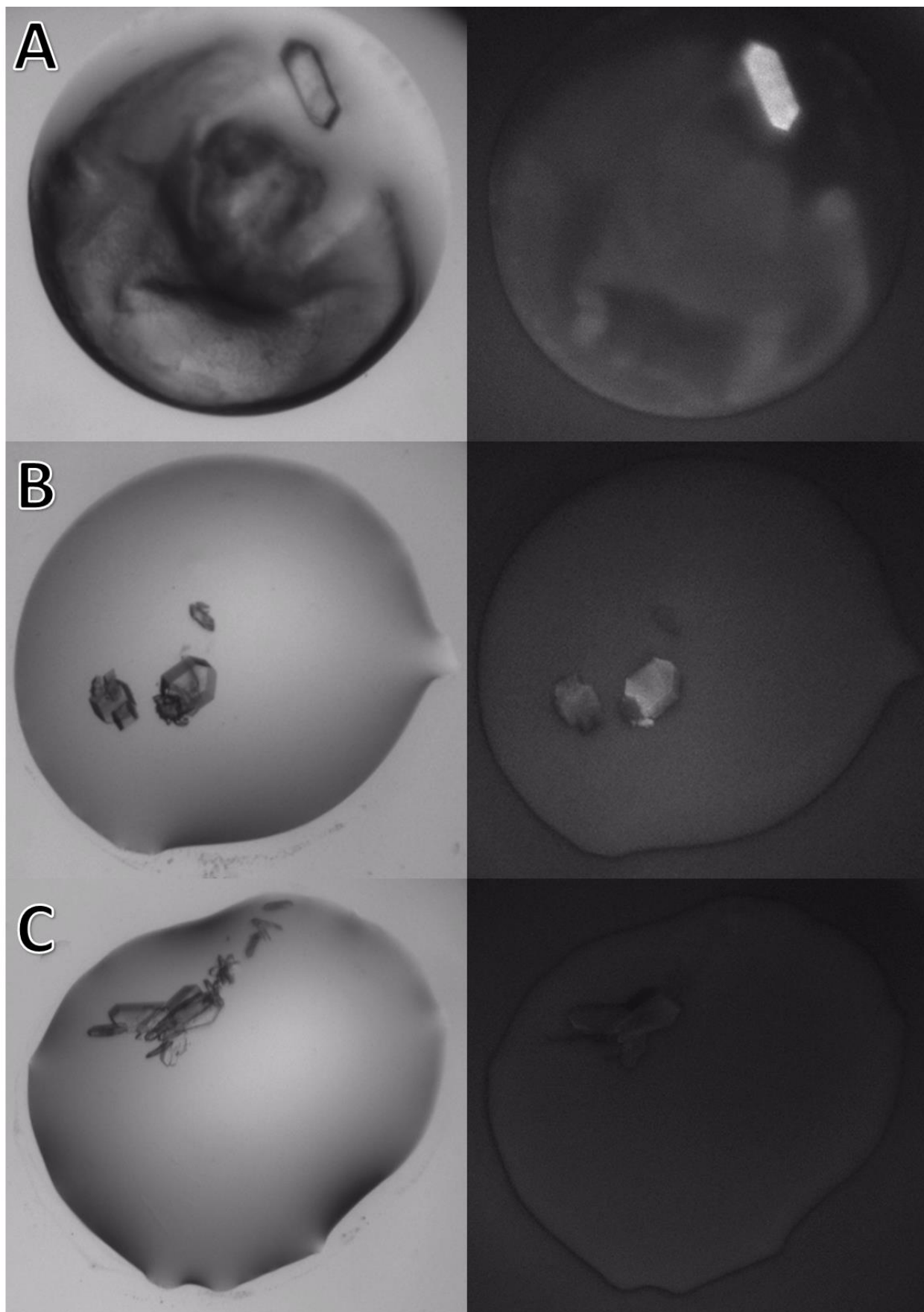
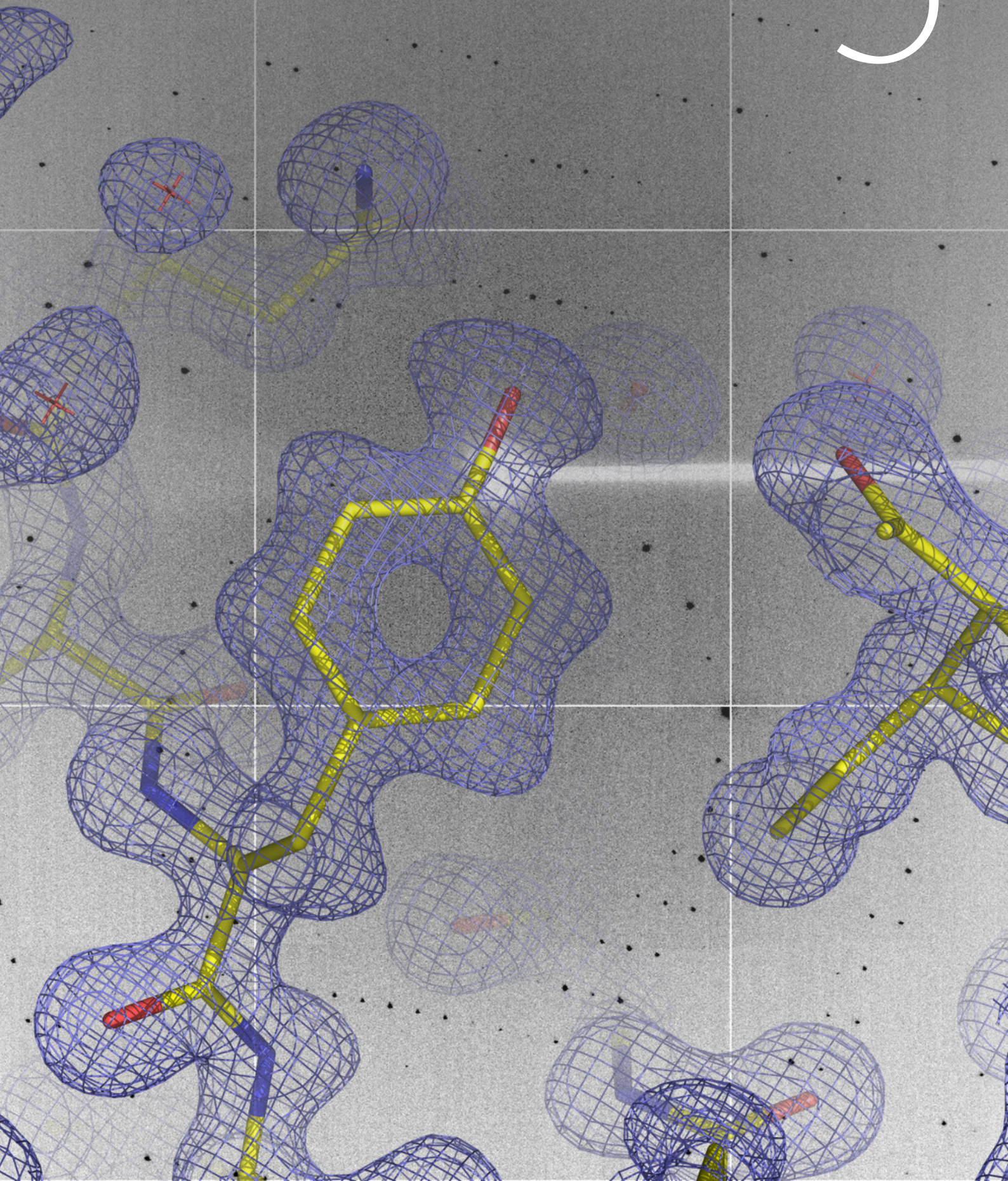


Figure S7. Visible (left) and UV (right) imaging of EcDHPS crystals as the (A) apo-enzyme, or bound to (B) **32** or (C) **46**. The intrinsic tryptophan fluorescence is quenched by the ligands and is an indicator of ligand occupancy.

Materials and Methods

5



5.1 Introduction

As chapters 2–4 of this thesis are directly reproduced from publications or prepared manuscripts some details of the experimental methods have been omitted. This chapter therefore contains additional information, explanation, and validation of the techniques used throughout this thesis, and the results therefrom.

The discovery of inhibitors targeting the folate pathway enzyme inhibitors HPPK and DHPS involved a variety of techniques commonly used in structural biology and SBDD. Proteins were expressed in *E. coli* bacterial expression systems and then purified using Fast Protein Liquid Chromatography (FPLC). Protein structure and protein-ligand interactions were investigated with methods including NMR spectroscopy, X-ray crystallography, SPR, enzymatic assays, and in silico methods, all of which are detailed in this chapter.

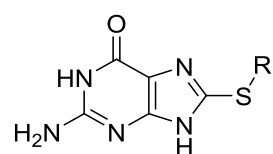
5.2 Compound procurement and synthesis

Compounds were purchased from suppliers as per Table 5.1. The inhibitors 8MG and compound **SAN3** were commercially sourced, however the other inhibitors were synthesised by Dr Aaron DeBono, Mr Noel Pitcher, Dr Michael Lee, Dr Zhong-Chang Wang, and Dr Jitendra Harjani. Inhibitors mentioned within this chapter are displayed in Table 5.2 for reference.

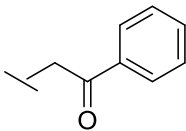
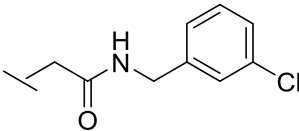
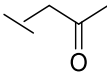
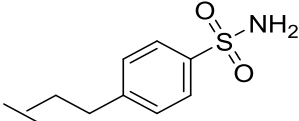
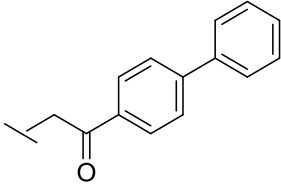
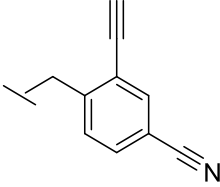
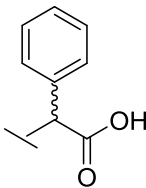
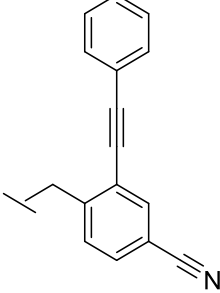
Table 5.1. Suppliers of reagents.

Supplier	Reagents
Sigma-Aldrich	Deuterium oxide ($^2\text{H}_2\text{O}$), ^{15}N -ammonium chloride ($^{15}\text{NH}_4\text{Cl}$), ^{13}C -glucose, dimethyl sulfoxide (DMSO), bovine serum albumin (BSA), kanamycin, sodium dodecyl sulfate (SDS), ammonium persulfate ($(\text{NH}_4)_2\text{S}_2\text{O}_8$), tetramethylethylenediamine (TEMED), tris(hydroxymethyl)aminomethane (tris), 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), dithiothreitol (DTT), citric acid trisodium salt, thiamine, adenosine triphosphate (ATP), α,β -methyleneadenosine 5-triphosphate (AMPCPP), biotin, 8-

	mercaptoguanine (8MG), thrombin, tryptone, agar, magnesium chloride, sodium chloride, deoxyribonuclease (DNase), β -mercaptoethanol
Bio-Rad Laboratories	Acrylamide/bis-acrylamide, Precision Plus Protein standard
Roche	EDTA-free Complete protease inhibitor tablets
TimTec	2-Amino-8-[[2-(4-methoxyphenyl)-2-oxoethyl]sulfanyl]-3,7-dihydro-6 <i>H</i> -purin-6-one (compound SAN3 , Table 5.2)
AMRESCO	Yeast, sodium phosphate monobasic, potassium phosphate dibasic
Cambridge Isotope Laboratories	Dimethyl sulfoxide- d_6 (DMSO- d_6), deuterium oxide (2H_2O)
Geneart	Plasmid vectors and genes
GE Healthcare	<i>N</i> -hydroxysuccinimide (NHS), <i>N</i> -ethyl- <i>N'</i> -(3-diethylaminopropyl)carbodiimide (EDC)
Astral Scientific	Isopropyl β -D-1-thiogalactopyranoside (IPTG), Coomassie Brilliant Blue
Merck Millipore	BugBuster Protein Extraction Reagent, ethanol, acetone, glycerol
Promega	KinaseGlo assay kit
Agilent	BL21 (DE3) competent <i>E. coli</i> cells
Schircks Laboratories	Pteric acid, 6-hydroxymethylpterin, 6-hydroxymethylpterin-monophosphate, 6-hydroxymethylpterin-diphosphate

Table 5.2. Chemical structures of inhibitors mentioned in this chapter

Compound	R	Compound	R
SAN3		8843	
8215		8921	
8220		8922	
8250		NP78	
8252		NP120	
8253		NP134	
8254		NP148	
8255		MLoF	
8409		MLOpF	

8411		ML-9-80	
8412		ML-9-100	
8683		JRH144	
8684		JRH162	

5.3 X-ray crystallography

Crystallisation trials were set up by the CSIRO Collaborative Crystallisation Centre (C3).¹ This facility uses automation and robot handling to prepare crystallisation plates. The TECAN EVO 100 (Tecan) and Phoenix liquid handling system (Art Robbins Industries) allow for both easy customisation of crystallisation conditions in addition to a reduced droplet size. The typical droplet composition was 150 nL of protein sample and 150 nL of reservoir or, alternatively, for trials that employed seeding,² 150 nL of protein sample, 120 nL of reservoir, and 30 nL of seed stock. A 50 μ L aliquot of protein is, in this manner, sufficient for 2 96-well plates. This sitting-drop vapour-diffusion method uses Innovaplate™ SD-2 plates (Innovadyne).

Seed stocks were prepared based on a published protocol.² A suitable drop (one with crystals and minimal precipitate) was chosen; 1 mL of the drop's reservoir solution was prepared. When using multiple drops as seed sources, the most concentrated reservoir was chosen to be prepared. A pinch of 400 μ m silica beads (OPS Diagnostics) was placed in a 1.5 mL Eppendorf tube and 50 μ L of reservoir solution added. Using a 2 μ L pipette, the

seed source droplet(s) was moved to the 50 μL of reservoir solution containing the seed beads. The mix was vortexed for one minute, an additional 200 μL of the reservoir solution was added, and then vortexed to mix. When preparing seed stock, the final dilution can be modified to control the degree of nucleation that occurs.

Plates were incubated at either 8 or 20 $^{\circ}\text{C}$. Crystal growth was monitored via visible light and UV light imaging. Drops were imaged 7 times in the first two weeks, and then weekly until 10 weeks. Drops were imaged with UV light at 1 week, 5 weeks, and 10 weeks.

To confirm that the crystals in question are protein, rather than salt or small molecule, a PX Scanner (Rigaku) was used to perform X-ray diffraction on crystals directly from crystallisation plates. The diffraction pattern of salt crystals is readily distinguishable from that of protein crystals, and undesired crystals are identified. As this instrument is run at room temperature, and crystals are often temperature sensitive (being observed to sometimes dissolve upon a temperature change), only crystals that were grown in the 20 $^{\circ}\text{C}$ incubator were tested in this manner. Crystals grown in the 8 $^{\circ}\text{C}$ incubator relied on UV imaging to assist in classifying each crystal.

Crystals were typically mounted using Dual-Thickness MicroLoopsTM (MiTeGen). To cryoprotect the crystal, cryoprotectant is added at the desired concentration to the reservoir, mixed, and then reservoir added to the drop. Typically this was glycerol, ethylene glycol, or a 50:50 mixture thereof. When using oil-based cryoprotectants, these were added directly to the drop, the crystal being passed through the oil prior to flash-freezing.

X-ray data were collected at the Australian Synchrotron MX-1³ or MX-2 beamlines using an ADSC Quantum 210 or 315 detector, respectively. Data were indexed using XDS⁴ and scaled using SCALA⁵ or Aimless.⁶ The initial phases were solved by molecular replacement using Phaser.⁷ Refinement was predominantly performed using REFMAC5,⁸ (some refinement performed using Phenix_refine⁹) and the electron density maps were visualised in Coot.¹⁰ After several rounds of manual rebuilding, ligands and water molecules were added, and the models further refined. PDB_REDO¹¹ was used to help inform ideal refinement parameters. Crystallographic information files (CIFs) for the ligands were generated using AFITT.¹² MolProbity¹³ and the wwPDB¹⁴ validation server were used to validate structures prior to deposition to the PDB.¹⁵

SaHPPK crystallisation – The concentration of SaHPPK used in crystallisation trials was 6.9–7.5 $\text{mg}\cdot\text{mL}^{-1}$ (~400 μM). A 2-fold molar excess of ligand, cofactor, and magnesium ions

over protein was deemed appropriate, based on the known affinities. One molecule of ligand and cofactor, and two magnesium ions, are present per SaHPPK molecule, and therefore ligand/AMPCPP and MgCl_2 were added to a final concentration of 1 mM and 2 mM, respectively. AMPCPP and MgCl_2 were prepared as 100 mM stocks in Milli-Q water, while compounds were typically prepared as 50 mM DMSO stocks. The total dilution was therefore typically 1.05-fold, with a DMSO concentration of ~2% [v/v], for ligand/AMPCPP/SaHPPK complexes. For the AMPCPP/SaHPPK complex the addition of AMPCPP and MgCl_2 diluted the protein samples 1.03-fold. The ligands used throughout this study were typically not soluble at 1 mM in aqueous solution, however, and precipitated upon addition to the protein sample. Samples were left on ice 1–2 h to allow for equilibration, before brief centrifugation (SpinFuge SF7000 (Bioline)) and transfer of the supernatant to a new PCR tube. This step helped prevent undissolved compound from blocking the robot dispensing needles.

Initially, a broad crystallisation screen consisting of eight 96-well plates (C3 Screens, CSIRO, Table 5.3 summarises the conditions used in each plate) was performed with AMPCPP/SaHPPK and **SAN3**/AMPCPP/SaHPPK. **SAN3**/AMPCPP/SaHPPK crystals grew after 3 days in a well containing 0.2 M ammonium fluoride and 20% [w/v] PEG 3350, and another well containing 0.2 M sodium nitrate and 20% [w/v] PEG 3350. A further crystallisation plate was set up using various concentrations of sodium nitrate, sodium fluoride, ammonium nitrate, or ammonium fluoride with PEG 3000, PEG 3350, or PEG 4000. The solved **SAN3**/AMPCPP/SaHPPK complex was obtained from a well containing 0.275 M ammonium nitrate and 22.1% [w/v] PEG 4000 (Figure 5.1). Solving this structure revealed strong density for a nitrate ion interacting with various loop 3 residues and the loop 1 residue I12 (Figure 5.2). These interactions likely stabilise these loop regions, and promote crystallisation, thus clarifying why the best crystal growth was observed in conditions containing nitrate.

Crystals of the AMPCPP/SaHPPK complex grew in conditions containing ~0.2 M salt, PEG, and tris-chloride. A plate screening different salts, and different concentrations of both tris-chloride and PEG 8000 was performed. Visually, the crystals obtained did not show much improvement (Figure 5.3), despite providing unexpectedly improved diffraction (2.4–2.7 Å). A dataset of 2.7 Å resolution possessed the best density for AMPCPP, this crystal being grown in 0.12 M magnesium acetate, 12.6% [w/v] PEG 8000, and 0.12 M tris-chloride, pH 8.5. Analogous to the nitrate molecule in the previous complex, the tris molecule is

likewise observable in the electron density, seen to form a stabilising hydrogen bond network between the SaHPPK monomers (Figure 5.2).

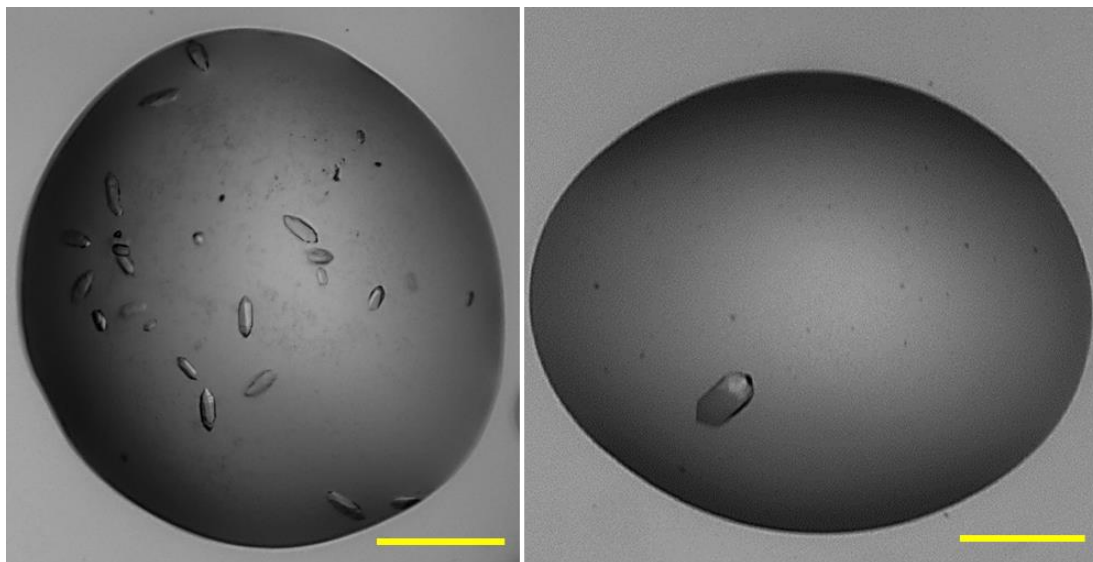


Figure 5.1. Initial (left) and final (right) crystals of the **SAN3/AMPCPP/SaHPPK** complex. All scale bars indicate 250 μm .

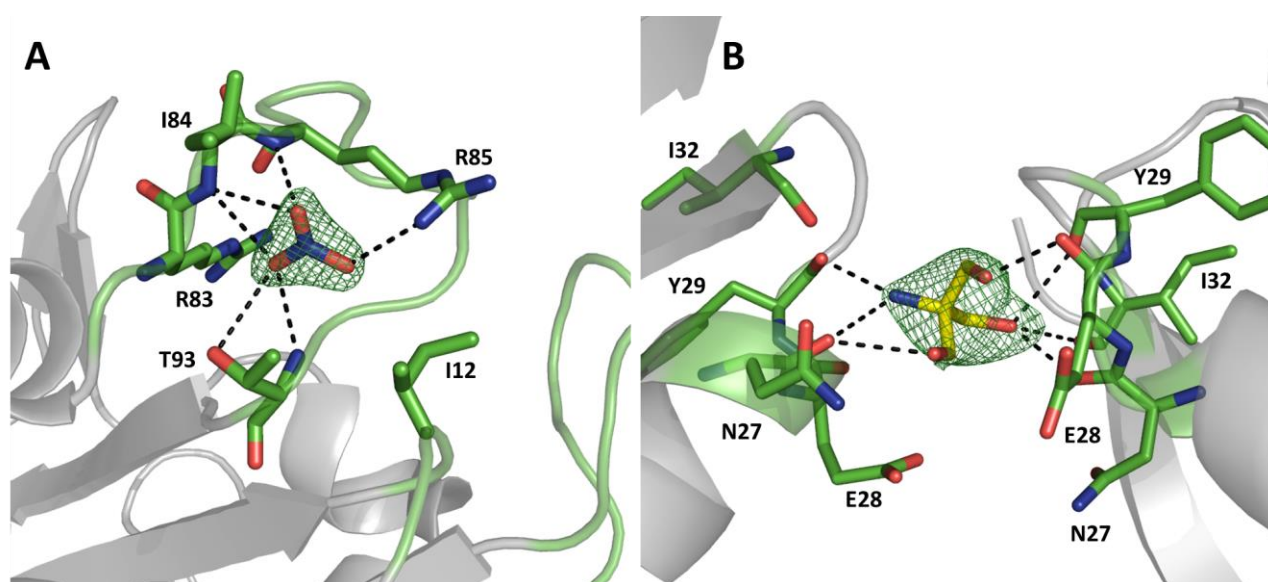


Figure 5.2. (A) Nitrate ion interactions in the **SAN3/AMPCPP/SaHPPK** crystal structure. (B) Tris interactions between SaHPPK monomers in the **AMPCPP/SaHPPK** crystal structure. Tris is coloured yellow. Hydrogen bonds are displayed as dashed black lines. The mFo-DFc difference density maps are contoured to 3σ .

Table 5.3. Description of ‘C3’ screens used during this thesis as reported on the C3 website (<http://c6.csiro.au/>).

Screen	Contents
c3_1	Peggy: Low MW, diverse buffers, pH range and salts.
c3_2	Peggy: Low-Mid MW, diverse buffers, pH range and salts.
c3_3	Peggy: 3350 and 4000, diverse buffers, pH range and salts.
c3_4	Peggy: Mid-High MW, diverse buffers, pH range and salts.
c3_5	Salty: Chloride, acetate, sulfate, formate and phosphate anions.
c3_6	Salty: Ammonium sulfate, lithium sulfate & trisodium citrate.
c3_7	Organics: MPD, 'small' branched polymers, hexanediol.
c3_8	Organics: Alcohols.

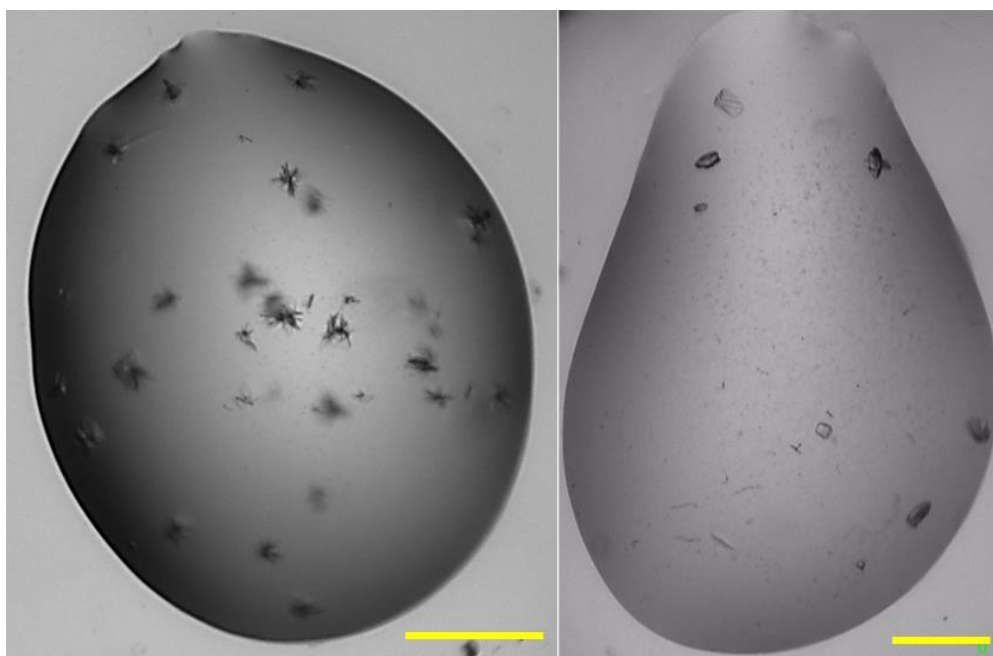


Figure 5.3. Initial (left) and final (right) crystals of the AMPCPP/SaHPPK complex. All scale bars indicate 250 μm .

The same focused screen used to obtain the **SAN3**/AMPCPP/SaHPPK complex was run using the compounds **8255** and **8252**, at both 8 and 20 °C. A lack of crystal growth was noted for both ligands indicating that the crystallisation of SaHPPK was ligand-specific. Compounds more similar to **SAN3**, namely **8411**, **8412**, and **8683**, were run using this screen — crystal growth was observed only for the **8683**/AMPCPP/SaHPPK complex, solved to 1.6 Å in a well containing 0.21 M ammonium nitrate and 22.2% [w/v] PEG 3350. For the **8411** and **8412** ternary complexes, only salt crystals were obtained. This was a surprising result particularly for **8411** considering the chemical similarity to **SAN3**, further illustrative as to the ligand-specificity of this crystal form. Compound **8255**, the lead SaHPPK compound at the time, was tested in the 8 C3 screens, a JCSG+ screen, and the above AMPCPP/SaHPPK focused screen to search for conditions that may allow crystal growth. Small crystal growth was observed in the presence of sodium acetate/tri-sodium acetate and PEG 8000/MPEG 5000 (Figure 5.4). A screen around these conditions was designed, compounds **8255** and **8409** being employed both with and without seeding, however there was little improvement over the initial crystals with only poor diffraction obtained (Figure 5.4).

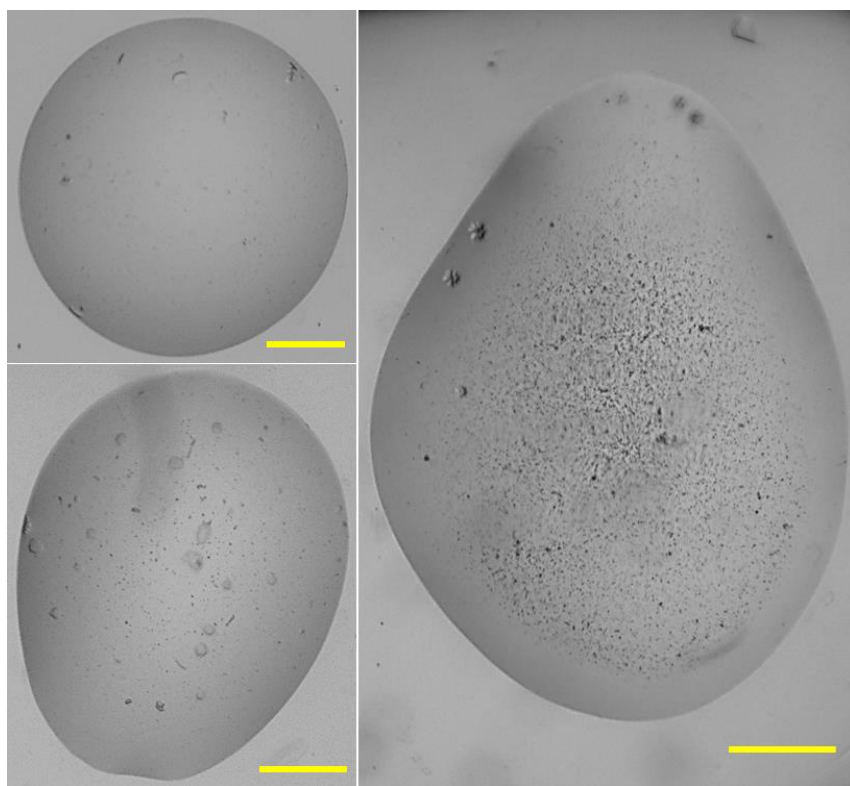


Figure 5.4. Initial (left) and final (right) crystals of the **8255**/AMPCPP/SaHPPK complex. Note that none of these crystals diffracted adequately for structure solution. All scale bars indicate 250 μm .

An additive screen (Hampton Research Additive Screen) was employed to broaden the chemicals being tested. Compound **8843** was chosen to test this strategy; of the 96 compounds tested, crystal growth was observed only in the presence of thiocyanate (0.05 M, Figure 5.5). A screen of thiocyanate in the range 0.025–0.8 M gave the best crystal growth at lower concentrations and 0.05 M was chosen as the ideal. A JCSG+ screen was employed with the addition of 0.05 M thiocyanate to optimise the other chemicals present. Although a **8843**/AMPCPP/SaHPPK complex remains unsolved, a high-diffracting crystal (2.0 Å) of the **8921**/AMPCPP/SaHPPK complex was in this way obtained from a reservoir solution containing 0.05 M thiocyanate, 20% [w/v] PEG 8000, 0.1 M tris-Cl, pH 8.5, and 0.2 M MgCl₂ (Figure 5.5).

All subsequent ligand/AMPCPP/SaHPPK structures were obtained from the above conditions, or the conditions from which the **SAN3**/AMPCPP/SaHPPK complex was obtained, dependent on whether they were benzyl- or acetophenone-functionalised, respectively. SaHPPK ternary complexes with **8254**, **8922**, **NP120**, and **NP134** (Chapter 3) were thus obtained in this manner. Growth and diffraction of both of these crystal forms is highly ligand-dependent; the currently unpublished AMPCPP/SaHPPK complexes of **JRH144** and **MLopF** were successfully acquired (Figure 5.6), however complexes with the similar ligands **JRH162** and **MLoF** appear intractable. In the case of **JRH162**/AMPCPP/SaHPPK, crystal growth was successful, however diffraction was never better than 4 Å. For the **MLoF**-bound complex, no crystal growth was observed. Attempts to crystallise several other ligands, including the DHPS lead compound, **8684**, likewise proved unsuccessful.

Crystallisation trials with HMDP/AMPCPP or HMDP-PP have been performed in the “shotgun” screen created by the C3 facility. This is a 1 plate broad screen designed to maximise the chances of successful crystal growth. To date, crystals of the HMDP/AMPCPP/SaHPPK have been observed in wells containing 0.1 M tris-chloride and PEG 3350, however these require optimisation before the corresponding structure is expected to be solved.

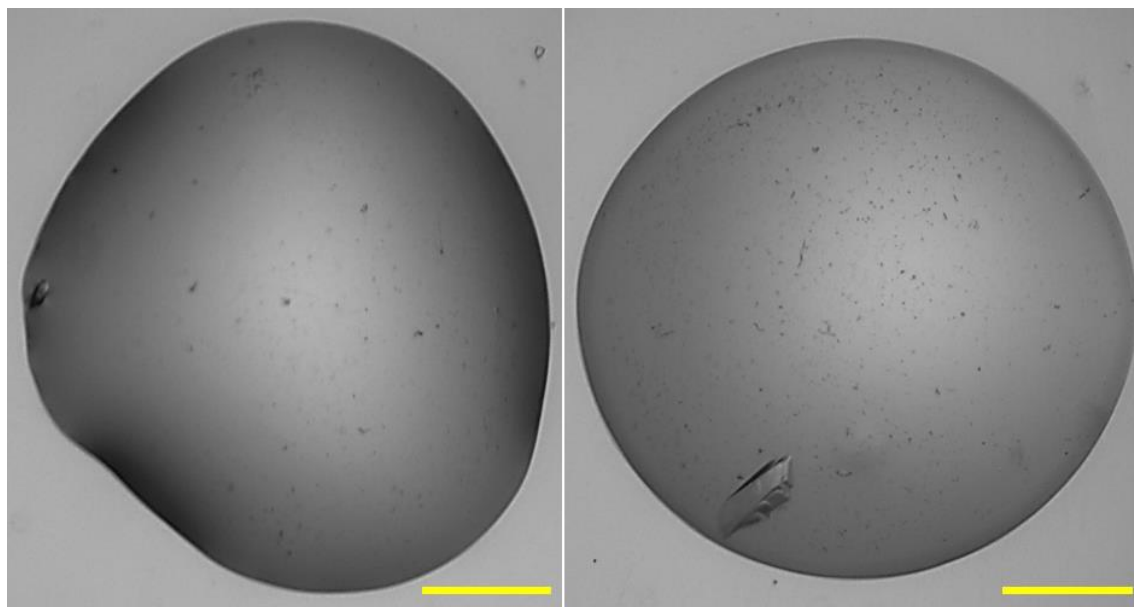


Figure 5.5. Initial crystal of **8843**/AMPCPP/SaHPPK (left). Final crystal of **8921**/AMPCPP/SaHPPK. All scale bars indicate 250 μm .

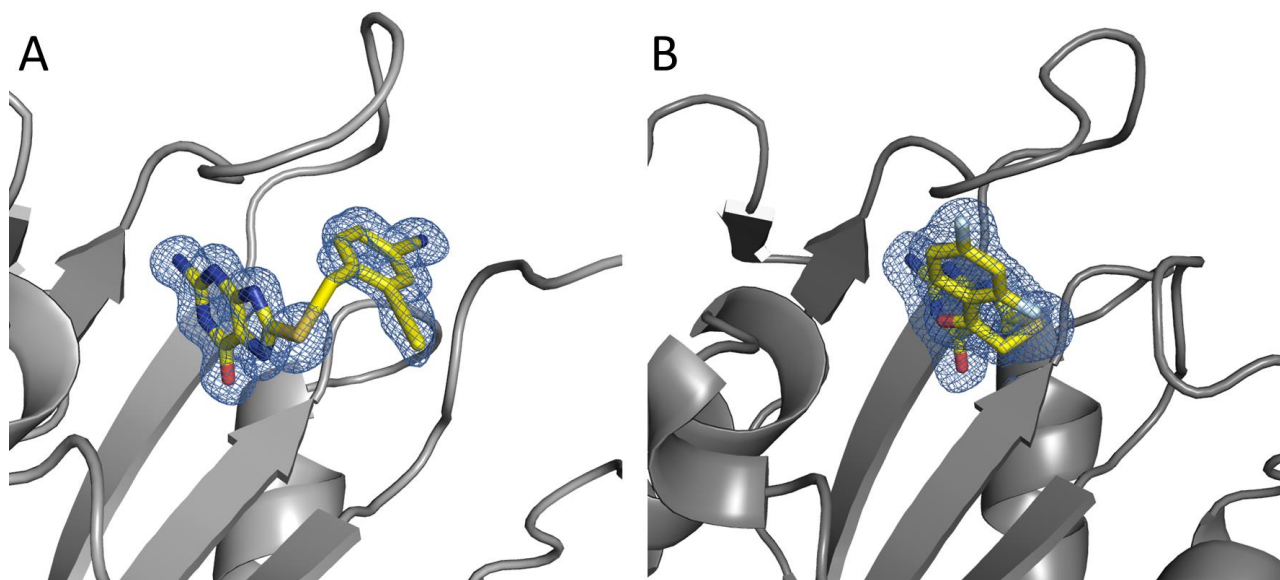


Figure 5.6. Crystal structures of AMPCPP/SaHPPK bound to (A) **JRH144** or (B) **MLoF**. $2F_o - F_c$ maps of the ligands are contoured to 1σ .

EcHPPK crystallisation – EcHPPK crystallisation was considerably more straightforward and robust than that of SaHPPK. Of the compounds for which attempts to co-crystallise with AMPCPP/EcHPPK were made, namely **8255**, **8250**, **8254**, **8922**, **NP134**, **NP120**, and **NP148**, only the **NP148**/AMPCPP/EcHPPK did not produce suitable crystals. In this case, well-diffracting crystals were obtained, however no ligand was observed to bind.

The concentration of EcHPPK used in crystallisation trials was 6.6 mg·mL⁻¹. As per SaHPPK, 1 mM of ligand and AMPCPP, and 2 mM of MgCl₂ was added to the protein solution. The conditions reported in ref 16 were used as a starting point for crystallisation trials. The **NP134** ternary complex was obtained in a reservoir solution of 20% [w/v] PEG 4000, 0.1 M tris-chloride, pH 9.1, and 0.172 M CaCl₂ (Figure 5.7). The complexes of **8255**, **8250**, **8254**, **8922**, and **NP120** were obtained from solutions containing approximately 0.1 M sodium-HEPES, pH 7.5, 0.2 M CaCl₂, and 25–30% [w/v] PEG 4000. Despite being the native metal for this protein, the magnesium ions present were displaced in the binding site by the excess of calcium used in the reservoir solution (as determined by the high to atomic resolution obtained).

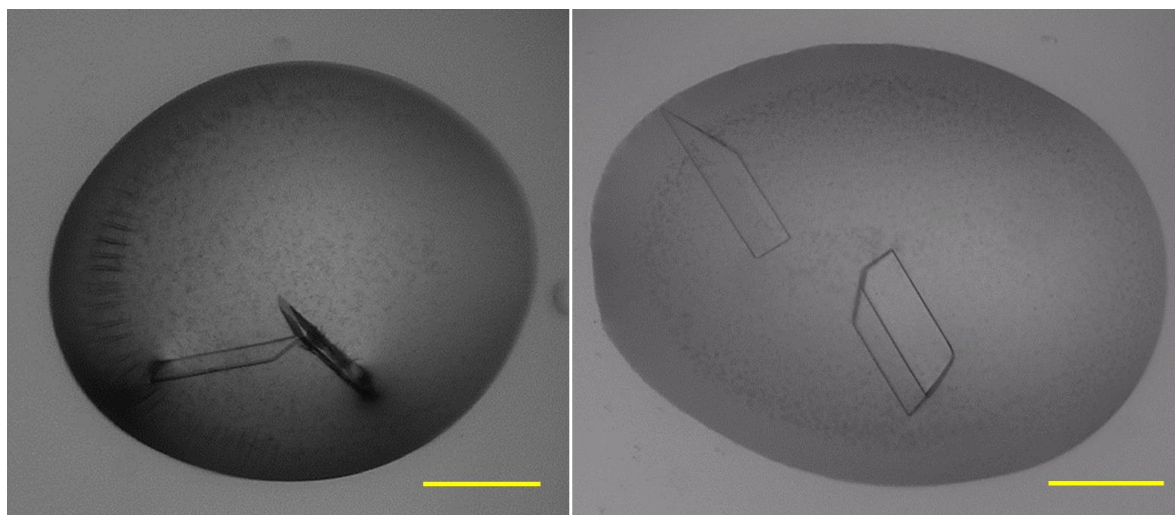


Figure 5.7. Crystals obtained for inhibitor/AMPCPP/EcHPPK complexes containing **NP134** (left) or **8255** (right). Structures of these complexes were solved to atomic resolution (1.05 and 1.09 Å respectively). All scale bars indicate 250 μm.

EcDHPS crystallisation – As per HPPK crystallisation, ligands were added to the DHPS protein solution at a concentration of 1 mM (the racemic compound **8684** added to a concentration of 2 mM). In the solved crystal structures, EcDHPS was present at concentration of 10–11.1 mg·mL⁻¹ (~350 µM). Initially, crystals were readily obtained, however these were thin needles, growing only in clusters, and poorly diffracting (~4–10 Å) (Figure 5.8, left). Many crystallisation screens that investigated factors including precipitants, protein concentration, temperature, drop size, drop ratio, and seeding were employed to improve the crystal quality. Thicker, better diffracting crystals were obtained (Figure 5.8, middle), however the diffraction was still no better than 3 Å, with poor ligand density. All EcDHPS crystal screens to this point were performed with a protein buffer of 50 mM tris-chloride, pH 8.5, 2 mM MgCl₂. A differential scanning fluorimetry (DSF) screen (vide infra) returned new protein buffer conditions (50 mM trisodium citrate, 150 mM NaH₂PO₄, and 350 mM K₂HPO₄, pH 6.9); implementation of this buffer with the previous reservoir conditions quickly returned crystals of high quality (Figure 5.8, right), typically diffracting ~1.8–2.0 Å. The cryoprotectant was important to this diffraction; a range of cryoprotectants were tested, ethylene glycol (30% [v/v]) giving the best results. Complexes of EcDHPS with **8684**, **8255**, **SAN3**, **8215**, **ML-9-100**, and pteric acid were obtained by co-crystallisation, with reservoir conditions containing 0.083–0.156 M magnesium acetate/sulfate/chloride, 17–28.6% [w/v] PEG 8000/PEG 6000/MPEG 5000, and 0.1 M tris-chloride or sodium cacodylate, pH 6.3–8.8. A **8220**/EcDHPS complex was obtained via soaking of an apo-crystal grown in 0.2 M MgCl₂, 25% [w/v] PEG 3350, and 0.1 M sodium-HEPES, pH 7.5. To soak the compound, 5 µL of a 10 mM DMSO stock was added to the reservoir (9.1% DMSO, 0.91 mM **8220** final concentration), the reservoir was then mixed before being added (2 µL) to the drop. The plate was resealed and left overnight before flash freezing of the crystal (AP/E core 150 oil (Mobil/Exxon)) as cryoprotectant). Magnesium phosphate crystals were routinely observed in the drops, however these were morphologically distinct from the desired protein crystals.

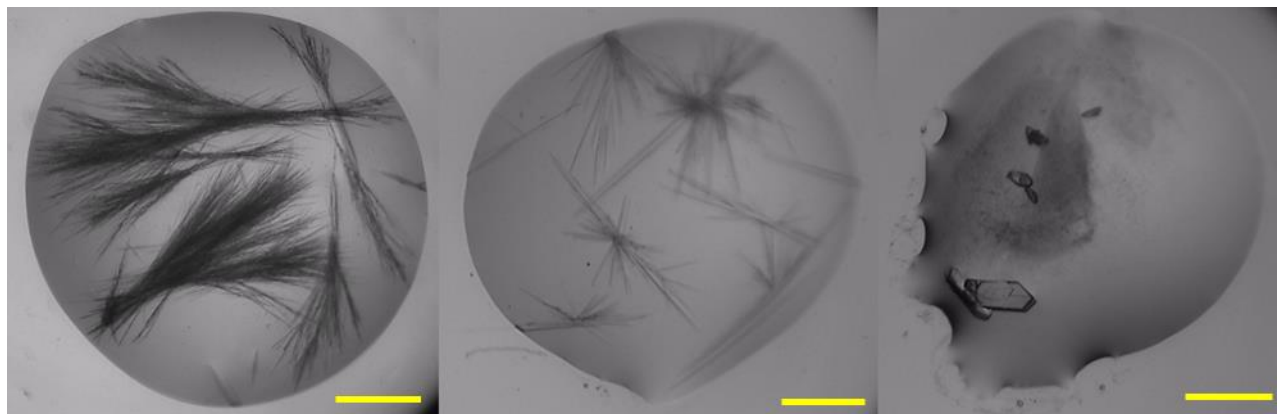


Figure 5.8. Crystals of **8684**/EcDHPS throughout optimisation (left to right). All scale bars indicate 250 μm .

5.4 Bacterial cell preparation

Escherichia coli is a common expression host and was used to produce all proteins studied in this thesis. pET28a vectors (encoding for kanamycin resistance) and containing the protein sequence of interest with a N-terminal hexahistidine tag and either a thrombin (SaHPPK and EcDHPS) or TEV (EcHPPK) cleavage site, were purchased from Genentech. The 5 ng plasmid samples were made up in 50 μL Milli-Q water and dispensed as 10 μL aliquots. Transformation was performed using a heat shock method. 50 μL of BL21 (DE3) cells (Agilent, $100 \text{ ng} \cdot \mu\text{L}^{-1}$) were thawed on ice for 30 min prior to the addition of 1 μL of plasmid. Cells remained on ice for a further 10 min prior to being placed in a 42°C water bath for 45 s. Cells were put back on ice for 2 min, 200 μL of LB was added, and samples were placed in a 37°C incubator at 200 rpm for 1 h. 200 μL of cell culture was then added onto an LB agar plate containing kanamycin to select for cells that had been successfully transformed with the plasmid. Plates were incubated at 37°C overnight and then stored at 4°C . Glycerol stocks were prepared by inoculating 5 mL of LB media with a single colony and growing at 37°C , 200 rpm, until an optical density of 600 nm (OD_{600}) ~ 0.6 was reached. Samples were diluted to a final glycerol concentration of 40% [v/v] and stored, in 1.5 mL Eppendorf tubes, at -80°C .

5.5 Protein expression

Protein expression and purification formed a significant component of the laboratory work throughout this thesis. The techniques employed in this regard are discussed below.

Ethanol (80% [v/v]) spray and a sterile fume hood (TopSafe, BioAir) was used to maintain an aseptic environment. Cultures were grown in a Multitron incubator (INFORS HT) in baffled flasks with shaking at ~200 rpm. The OD₆₀₀ of a culture was used to monitor its growth, measured using a Cary® 50 UV-Vis spectrophotometer (Agilent). When expressing protein, an agar plate containing the appropriate antibiotic (kanamycin for the three proteins presented in this thesis) was streaked from a glycerol stock, prior to inoculation of a culture (EcDHPS expression particularly required this step as inoculation direct from a glycerol stock would typically show markedly reduced expression levels). Typically, 5 mL of either rich or minimal media culture was inoculated with a single colony. These cultures were grown for ~4 h at 37 °C or, when left overnight, at 30 °C to prevent excessive growth. This culture was then passaged, grown, and expressed using the methods described below. These cultures all contained kanamycin at 50 µg·mL⁻¹. Following expression, cultures were typically centrifuged at 3800 × *g* using a JA-14 rotor in an Avanti® J-25 centrifuge (Beckman) for 15 min, and then cell pellets stored at -80 °C. Protein expression levels and protein purity were assessed by SDS-PAGE. SDS-PAGE sample buffer was added and samples heated to 90 °C, prior to loading onto the prepared gels. When ascertaining soluble vs insoluble expression, BugBuster® Protein Extraction Reagent (Merck Millipore) was added to samples; after 30 min, supernatant and pellet were separated with subsequent addition of SDS-PAGE sample buffer to each.

Media recipes are found in Appendix I. All ingredients were autoclaved or sterile filtered where appropriate. While LB was often used for overnight cultures, 2x YT media is associated with better cell density and protein yield,¹⁷ and was typically used in the expression of unlabelled proteins.

Autoinduction media was prepared as per Studier's seminal paper.¹⁸ This includes both a rich (ZYM-5052) and minimal (N-5052) autoinduction media, suitable for production of unlabelled and ¹⁵N-labelled proteins respectively.

¹⁵N¹³C-labelled protein was expressed using M9 minimal media¹⁹ in which a grown overnight culture is centrifuged (5 min, 3200 × *g*, 30 °C) in an Eppendorf™ Model 5810 centrifuge (Fisher Scientific) and resuspended in 50 mL of (pre-warmed) unlabelled M9 minimal media. This culture is grown at 37 °C until mid-log phase (OD₆₀₀ ~0.5–0.8), centrifuged (20 min, 3200 × *g*, 30 °C), then resuspended in 500 mL of M9 minimal media with labelled or unlabelled NH₄Cl and glucose as desired. At mid-log phase, the temperature is reduced and the culture induced with a determined amount of IPTG (typically 0.5 mM).

Unlabelled expression was performed by passaging an overnight culture into fresh 2x YT media and growing at 37 °C until mid-log phase, before reducing the temperature and inducing with IPTG.

A colony selection protocol was undertaken to optimise yield as published.¹⁷ In this method, transformed cells are streaked onto agar plates and incubated. Four colonies are streaked onto each quarter of a new plate, then used to inoculate 2 mL cultures. These cultures are grown at 37 °C to an OD₆₀₀ of 2–3, then centrifuged at 1500 × *g* for 5 min (30 °C). Cells are resuspended in 5 mL of media; after 1 h cultures were induced with 0.3 mM IPTG at 20 °C for ~16 h. Samples are tested via SDS-PAGE; the colony displaying the highest expression level is chosen for further selection, and the process is repeated. The final colony is then prepared into a glycerol stock and stored at -80 °C. This “double selection” process has been demonstrated to improve the quantity and reliability of protein expression.¹⁷ Note only cells expressing SaHPPK or EcDHPS proteins underwent this “double selection” process.

“Colony-optimised” cells were used to test ideal IPTG levels. Cultures were prepared as above. A single culture (at mid-log) was split into 5 separate flasks and induced with 0.1, 0.3, 0.5, 0.7, or 1.0 mM IPTG at 20 °C for ~16 h. Samples were tested via SDS-PAGE. A maximum expression level was observed at an IPTG concentration of 0.5 mM for SaHPPK, EcHPPK, and EcDHPS.

Both IPTG induction and autoinduction expression protocols were trialled with SaHPPK to maximise yield. In the autoinduction protocol, an overnight culture is passaged into rich or minimal autoinduction media. Cultures are grown in this media for 6 h at 37 °C, the temperature is then lowered to 25 °C with expression for 24 h. ~7–14 g of cell paste was obtained per L for both rich and minimal media. Following purification, ~5 mg of SaHPPK was obtained per litre of culture. An equivalent rich autoinduction media with 2x YT instead of LB provided minimal improvement to this yield. Using the above-described M9 minimal media (expression at 20 °C, 0.5 mM IPTG) protocol, ~10 mg of SaHPPK was obtained per litre of culture. This protocol was thus used in subsequent expressions, due to both the improved yield relative to autoinduction, and the reduced quantity of labelled ingredients used (1.5 g vs 2.5 g of ¹⁵NH₄Cl).

Although some variation existed between runs, unlabelled-protein expression using IPTG induction produced ~10 mg, ~9 mg, and ~5 mg per litre of 2x YT culture, for SaHPPK,

EcDHPS, and EcHPPK, respectively. The temperature for protein expression was typically 25, 28, and 25 °C, for SaHPPK, EcDHPS, and EcHPPK, respectively.

5.6 Purification

Typically purification involved lysis of the cells using both lysozyme and sonication (XL2000, Misonix), centrifugation, a Ni-NTA IMAC column step, with a size exclusion column as a final buffer exchange and purification step. Column purification steps were performed using an ÄKTA purifier (GE Healthcare).

Cell pellets were lysed in 10 mL of buffer per gram of cell paste with addition of an EDTA-free Complete protease-inhibitor cocktail tablet (Roche) and lysozyme to a final concentration of 0.2 mg·mL⁻¹ (i.e. 2 mg per gram of cell paste). The solution was stirred at 4 °C for 15–60 min before sonication (10 rounds of 15/45 sec on/off). The lysate was centrifuged for 30 min at 39,000 × *g* using a JA-25.5 rotor in an Avanti® J-25 centrifuge (Beckman). The supernatant was filtered (0.45 µm filter) and loaded onto a Ni-NTA IMAC Column (Qiagen). Unbound protein was washed off with a 10 mM imidazole-containing buffer. Protein was eluted from the column with a high concentration of imidazole. The protein sample was loaded onto either a 16/60 or 26/60 Superdex 75 size-exclusion column (GE Healthcare). Protein collected from this column was concentrated using a 3 kDa molecular-weight cutoff ultrafiltration centrifugal device (Amicon) and snap frozen using liquid nitrogen or acetone/CO₂. Samples were tested via SDS-PAGE to confirm purity. Throughout the procedure, fresh DTT (1 mM) is added after each step. Protein concentration was determined using an ND-1000 spectrophotometer (Nanodrop).

All proteins were cloned with an N-terminal hexahistidine tag and either a TEV (EcHPPK) or thrombin (SaHPPK, EcDHPS) cleavage site. For crystallographic studies, this tag was removed to promote crystallisation. Following elution from the Ni-NTA IMAC column, the buffer was exchanged to remove imidazole. The appropriate protease was added — for the HPPK proteins, to promote stability, MgCl₂, ATP, and a ligand were also added — and the sample gently mixed overnight at 4 °C for 16 h. The sample was reloaded onto the Ni-NTA IMAC column and collected, with subsequent purification using a size exclusion column as above.

SaHPPK was lysed in a buffer containing 50 mM HEPES, 5% [v/v] glycerol, pH 8.0. Unbound protein was washed off the Ni-NTA IMAC column with 10 mM imidazole in 50 mM

HEPES, 0.3 M NaCl, 5% [v/v] glycerol, pH 8.0. Tagged protein was eluted in the buffer containing 250 mM imidazole. The buffer used during the size exclusion column step contained 50 mM HEPES, 1% [w/v] sorbitol, pH 8.0 (N.B. samples for X-ray crystallographic studies did not contain sorbitol).

EcHPPK was lysed in a buffer containing 50 mM HEPES, 5% [v/v] glycerol, pH 8.0. Unbound protein was washed off the Ni-NTA IMAC column with 10 mM imidazole in 50 mM HEPES, 0.15 M NaCl, 5% [v/v] glycerol. Tagged protein was eluted in a 500 mM imidazole-containing variant of this buffer. The buffer used during the size exclusion column step is that reported in ref 16, namely 20 mM tris-chloride, pH 8.0, 100 mM NaCl.

EcDHPS was lysed in a buffer containing 50 mM tris, pH 8.5, 5% [v/v] glycerol, 5 mM MgCl₂. Unbound protein was washed off the Ni-NTA IMAC column with 10 mM imidazole in 50 mM tris buffer, pH 8.5, 0.1 M NaCl, 5% [v/v] glycerol, 2 mM MgCl₂. Tagged protein was eluted in the same buffer with 500 mM imidazole. In initial purification runs, a MonoQ ion-exchange column (GE Healthcare) was used as a purification step. The column was equilibrated with 50 mM tris, pH 8.5, 2 mM MgCl₂, and protein eluted using the same buffer with the addition of 0.25 M NaCl. In later runs, it was found that using a 1 mL (rather than the previous 5 mL) Ni-NTA IMAC column improved the purity such that the ion-exchange step wasn't required. The buffer used during the size exclusion column step contained 50 mM tris, pH 8.5, 2 mM MgCl₂ or, alternatively, 50 mM trisodium citrate, 150 mM NaH₂PO₄, 350 mM K₂HPO₄, pH 6.9.

5.7 Surface plasmon resonance

Prior to SPR experiments, compounds were prepared at a range of concentrations in SPR running buffer. These were left overnight before being centrifuged and checked for precipitate. In this way an estimate of the maximum solubility of each compound was known prior to planning experiments.

SPR running buffer was prepared with the addition of all ingredients except DMSO. A small sample of DMSO-free buffer was set aside and then DMSO added to the desired final concentration. When preparing analyte samples for injection, this DMSO-free buffer can be mixed with the running buffer. In the correct ratio, this reduces the DMSO deviation of the sample to the running buffer and more accurate results are obtained. This DMSO-free buffer

is also used in preparing samples for solvent correction. All buffers are filtered prior to use in the SPR instruments.

Compounds were generally prepared in 10–50 mM DMSO stocks prior to being diluted into SPR running buffer. All samples are centrifuged at $18,000 \times g$ for 5 min (Microfuge® 22R, Beckman Coulter) before being plated. During the second half of this thesis, the VIAFLO ASSIST (INTEGRA) was used to dispense and dilute samples into plates for testing.

The SPR experiments were performed on two different instruments throughout this thesis, dependent on availability, namely a Biacore T200 biosensor (GE Healthcare) or a ProteOn XPR36 biosensor (Bio-Rad).

Biacore T200 biosensor – Preconditioning of an NTA chip (GE Healthcare) was performed prior to ligand immobilisation with 350 mM EDTA, pH 8.3, 3x 60 s injection $10 \mu\text{L} \cdot \text{min}^{-1}$. Immobilisations were performed in HBS-EP+ running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 50 μM EDTA, 0.05% [v/v] Tween-20) at 25 °C with a constant flow-rate of $10 \mu\text{L} \cdot \text{min}^{-1}$.

SaHPPK and EcDHPS proteins were covalently coupled to the NTA chip (GE Healthcare). This was performed using a method detailed in ref 20. Briefly, a single flow cell on the chip surface was sequentially activated by injecting (1) 40 μL of nickel sulfate and (2) 70 μL of a 1:1 mixture of NHS/EDC (*N*-hydroxysuccinimide/*N*-ethyl-*N'*-(3-diethylaminopropyl)carbodiimide). Recombinant protein was diluted in the running buffer (SaHPPK to $225 \mu\text{g} \cdot \text{mL}^{-1}$; EcDHPS to $80 \mu\text{g} \cdot \text{mL}^{-1}$) and injected over an activated flow cell for 20 min (200 μL). The amine-coupled surface was subsequently blocked with 70 μL of 1 M ethanolamine, pH 8.0, and then further regenerated with two 10 μL injections of 350 mM EDTA prepared in running buffer. Using this coupling approach, average immobilisation levels achieved were 5400 RU for SaHPPK and 7200 RU for EcDHPS. Additionally, ubiquitin-specific-processing protease 7 (USP7) was coupled in a similar fashion to provide for an unrelated negative control surface (6600 RU). All SPR binding experiments were performed at 20 °C in SPR binding buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 5 mM DTT, 10 mM MgCl_2 , 0.05% [v/v] Tween-20, 5% [v/v] DMSO). Analytes were serially diluted (3-fold) in SPR binding buffer, injected for 30 s contact time at $60 \mu\text{L} \cdot \text{min}^{-1}$, and then allowed to dissociate for 60 s. Each analyte titration was performed in duplicate or greater. A control compound (initially 8MG in this project, but later **8255**) with a known affinity was intermittently tested to monitor the surface. Binding sensorgrams were processed, solvent-corrected, and double-referenced using Scrubber software (BioLogic Software, Australia).

SPR binding analysis of several of the compounds investigated in this thesis revealed dissociation rates that were not sufficiently slow to allow global fitting to a kinetic binding model, for which the k_d (dissociation rate constant) must typically be $<0.5\text{ s}^{-1}$ for SPR instruments to be able to capture sufficient data points during the dissociation phase. Therefore, to determine binding affinities (K_D values), responses at equilibrium for each analyte were fitted to a 1:1 steady-state affinity model available within Scrubber using 8MG as a reference (this is described in more detail in ref 21).

ProteON XPR36 biosensor – Preconditioning of a GLH (Bio-Rad) chip was performed prior to ligand immobilisation with 100 mM HCl, 50 mM NaOH, and 0.5% SDS (first injected horizontally, then vertically) for 60 s, $30\text{ }\mu\text{L}\cdot\text{min}^{-1}$. Biotinylated SaHPPK and EcHPPK were captured via streptavidin, while EcDHPS was direct-coupled. Streptavidin was coupled to the chip surface at 37 °C in HBS-EP+ running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 50 μM EDTA, 0.05% [v/v] Tween-20) after activation with a 1:1 mixture of NHS/EDC. Protein was diluted to $100\text{ }\mu\text{g}\cdot\text{mL}^{-1}$ at pH 4 and injected for 6 min. The surface was then blocked with 1 M ethanolamine, pH 8.0, with an immobilisation of 15,000–16,000 RU for each flow cell.

EcDHPS was coupled to the surface at 25 °C in a running buffer containing 50 mM HEPES, 150 mM NaCl, 10 mM MgCl_2 , 3% [v/v] DMSO, 0.05% [v/v] Tween-20, pH 8.0, after activation with a 1:1 mixture of NHS/EDC. Protein was diluted to $\sim 200\text{ }\mu\text{g}\cdot\text{mL}^{-1}$ at pH 5.25 and injected for 6 min. The appropriate pH and sample dilution was determined in a pH scan (Figure 5.9). 6-methylamino-5-nitrosoisocytosine (a known pterin-site binder) was present at 1 μM in the injection mixture to protect the pterin-binding site. The surface was then blocked with 1 M ethanolamine, pH 8.0, with an immobilisation of $\sim 17,000$ RU. GRP-78 (78 kDa glucose-regulated protein) was immobilised as a control protein; protein was diluted to $58\text{ }\mu\text{g}\cdot\text{mL}^{-1}$ at pH 5.0 and injected for 5 min. The surface was then blocked as above with immobilisation levels of $\sim 16,400$ RU. EcHPPK and SaHPPK were minimally biotinylated using NHS-activated biotin and maleimide-activated biotin, respectively. HPPK immobilisation was performed at 20 °C in the same buffer used for EcDHPS. 1 mM of ATP was added to HPPK samples which were then injected over streptavidin-coupled flow-cells for 25 min. Immobilisation levels for EcHPPK and SaHPPK were $\sim 4,900$ and $\sim 3,800$ RU, respectively. Biotin was subsequently injected over HPPK-bound flow-cells to block remaining unbound streptavidin.

Analytes were serially diluted (3-fold) in SPR binding buffer and injected for 60 s contact time at $30 \mu\text{L} \cdot \text{min}^{-1}$, then allowed to dissociate for 60 s. Each analyte titration was performed in duplicate or greater. As above, a control compound was tested intermittently to monitor the surface. Binding sensorgrams were processed, solvent-corrected, and double-referenced using either the ProteOn Manager software, version 3.1.0 (Bio-Rad), or Scrubber (BioLogic Software, Australia). Compounds were tested from a top concentration of 3–300 μM , ideally ~ 10 times the K_D . Where solubility or affinity was limiting in this regard, compounds were globally fitted with a control compound that was approaching saturation to the protein.

The ProteOn biosensor used has a much reduced sensitivity compared to the Biacore biosensor (Figure 5.10). For this reason, a more complete sampling of concentrations (above and below the K_D) was generally required to acquire an accurate estimate of affinity. Many compound affinities that were determined on the Biacore biosensor were, in contrast, at a top concentration approximate to or below the K_D . This was not possible on the ProteON and, as such, when compounds displayed differing affinities to the proteins of interest, compounds were rerun with a concentration range appropriate to each enzyme.

Biotinylation – Lowering the pH of a protein's solution below its pI affords the protein an overall positive charge, which is then attracted to the negatively charged carboxylated dextran matrix of the chip. Unlike EcDHPS, which is stable at a pH below its pI, HPPK (in particular SaHPPK) is sensitive to changes in pH. This was overcome on the Biacore T200 through the use of an NTA chip, whereby the hexahistidine tag of HPPK is attracted to the surface, and the enzymes can therefore be captured (and then coupled) without perturbing the pH. An NTA chip (Bio-Rad) is also available for the ProteOn XPR36 biosensor; experiments on the HPPK enzymes using this chip were noted, however, to give inferior data quality. For this reason, the HPPK enzymes were biotinylated — SaHPPK via exposed cysteine residues and EcHPPK via exposed lysine residues. Streptavidin can then be directly coupled to the chip, with the biotinylated HPPK enzymes then captured at a pH better tolerated. EZ-Link Maleimide-PEG₂-biotin (Fisher Scientific) was added to SaHPPK at a 1:1 molar ratio and EZ-Link-NHS-LC-LC-Biotin (Fisher Scientific) was added to EcHPPK at a ratio of 0.8:1. SaHPPK and EcHPPK enzymes were at $0.63 \text{ mg} \cdot \text{mL}^{-1}$ and $0.84 \text{ mg} \cdot \text{mL}^{-1}$, respectively in a buffer containing 50 mM HEPES, 1% [w/v] sorbitol, 2 mM DTT, pH 8.0. Samples were left on ice for ~ 2 h to allow for the reaction to take place. Samples were then run on a Superdex 75 10/300 size exclusion column to remove unreacted biotin; HPPK samples were pooled across runs, dispensed as 50 μL aliquots, and snap frozen.

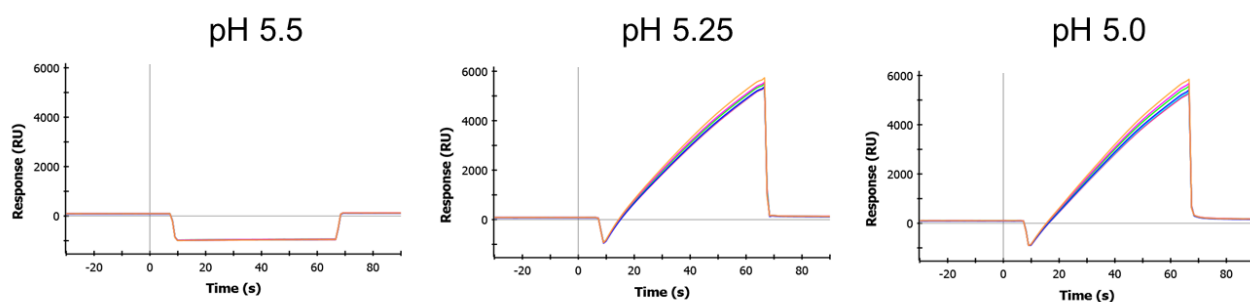


Figure 5.9. A pH scan of EcDHPS on a GLH chip (Biacore). Protein samples are diluted with buffers of varying pH, and injected over an unactivated surface. Only by dropping the pH below the pI of the protein will it attract to the chip surface. A 10-fold dilution of EcDHPS samples (final concentration $200 \mu\text{g}\cdot\text{mL}^{-1}$) to pH 5.25 showed the maximal attraction, with minimal aggregation on the surface (cf. pH 5.0 test which is starting to show tailing after the injection has finished).

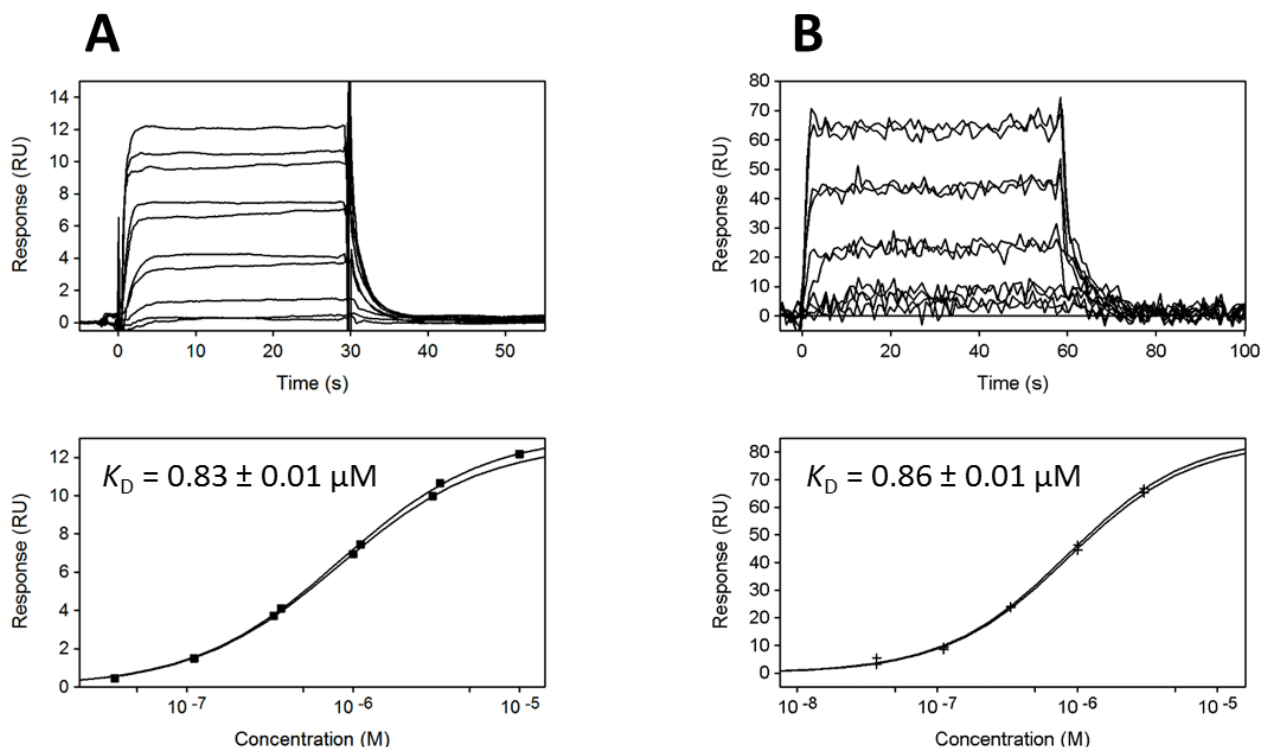


Figure 5.10. SPR sensorgrams (top) and steady-state binding curves (bottom) for compound **8220** when tested using the (A) Biacore T200 or (B) ProteOn XPR36.

5.8 NMR spectroscopy

All NMR experiments were recorded at 22 °C on a Bruker Avance 600 MHz NMR spectrometer equipped with a cryoprobe and Z axis gradient. Triple resonance NMR experiments were performed on a sample of ~0.25 mM $^{15}\text{N}/^{13}\text{C}$ -labelled SaHPPK dissolved in a 90%/10% $\text{H}_2\text{O}/\text{D}_2\text{O}$ 50 mM HEPES buffer, 1% [w/v] sorbitol, and 2% [v/v] $\text{DMSO}-d_6$ at pH 8.0 in the presence of 10 mM MgCl_2 and saturating amounts of AMPCPP (0.5–1 mM). Titrations of compounds **5720**, **8255**, **SAN3**, **8921**, **8253**, or **NP120** to saturation were performed from a 50 mM stock dissolved in $\text{DMSO}-d_6$. Backbone assignments were obtained using the HNCA experiment. Assignments for the **8255** and **SAN3** ternary complexes were further confirmed using HN(CO)CA, HNCO, and a ^{15}N -edited 3D NOESY experiment recorded with a mixing time of 120 ms. 3D experiments used a WATERGATE sequence for solvent suppression. ^{15}N heteronuclear NOE spectra were recorded on a ~0.36 mM ^{15}N -labelled sample of SaHPPK in the presence of 1 mM AMPCPP and either 600 μM 8MG or ~400 μM **SAN3** using gradients for coherence selection and sensitivity enhancement. Three seconds of saturation was applied using a binomial train of pulses separated by a delay of 5 ms to generate the desired heteronuclear NOE and was applied off- and on-resonance in an interleaved manner, in addition to 1 s of relaxation delay. Errors were calculated from the baseplane noise level. Spectra were processed using NMRPipe²² and analysed with XEASY²³ or SPARKY.²⁴ 2D ^{15}N FAST-HSQC experiments were typically acquired with $t_{1\text{max}}(^{15}\text{N}) = 51\text{--}62$ ms and $t_{2\text{max}}(^1\text{H}) = 142$ ms, whereas triple resonance experiments were acquired with $t_{1\text{max}}(^{15}\text{N}) = 23.3$ ms, $t_{2\text{max}}(^{13}\text{C}) = 10.4$ ms, $t_{2\text{max}}(^1\text{H}) = 15.1$ ms, and $t_{3\text{max}}(^1\text{H}) = 142$ ms. Experiments were performed in Shigemi NMR Tubes (www.shigeminmr.com).

5.9 Molecular modelling

Molecular modelling was performed using the Schrödinger Suite (www.schrodinger.com) through the Maestro interface (Maestro, Schrödinger, LLC, New York).²⁵ Protein preparation was performed with the Protein Preparation Wizard workflow implemented by Schrödinger (Epik, Schrödinger, LLC, New York),²⁶ with deletion of all waters. In order to eliminate any bond length or bond angle biases in the structures, compounds were subjected to a full minimisation prior to docking using LigPrep (LigPrep, Schrödinger, LLC, New York).²⁷ Docking was carried out with Glide.²⁸

To promote poses that conform to experimental data, positional constraints were frequently implemented when docking, for atoms of the 8MG moiety — poses of 8MG-derivatives would need to match the approximate (<0.5 Å) positions of 8MG as observed in the solved crystal structures. This practice helped to give more reliable binding poses, and was thus used when informing inhibitor design.

5.10 Isothermal titration calorimetry

ITC experiments were performed using an iTC200 instrument (MicroCal) at 25 °C, with ligands titrated into solutions of SaHPPK using 19x 2.2 μ L or 13x 3.1 μ L injections. Data were fitted using Origin software to yield the thermodynamic parameters, ΔH , K_d and N (the binding stoichiometry), assuming a cell volume of 0.2 mL. These were then used to calculate the Gibb's free energy of binding (ΔG) and entropy of binding (ΔS).

A solution of SaHPPK was exchanged into ITC buffer (50 mM HEPES, 1 mM TCEP, 10 mM $MgCl_2$, pH 8.0) using a Zeba Spin Desalting column (Thermo Fisher Scientific).

Titration with ATP were performed to compare results with previous literature values.²⁹ For these titrations with ATP, SaHPPK and ATP were at 70 μ M and 1500 μ M, respectively. For titrations with inhibitors, SaHPPK and inhibitor were at 15 μ M and 125 μ M, respectively, while ATP was present at 1.5 mM in both the cell and syringe. Inhibitors were initially prepared as a 50 mM stock solution in DMSO, and diluted 400-fold; an equal 0.25% [v/v] DMSO was thus added to the SaHPPK sample to ensure buffer matching. Experiments were limited by the solubility of 8MG, estimated to be ~ 300 μ M from ITC experiments assuming a 1:1 binding stoichiometry. Controls consisted of a titration of buffer into buffer, buffer into protein, and ligand into buffer.

While ITC experiments provided thermodynamic binding information for the few compounds tested, the consumption of protein was approximately 250 μ g for each compound (when tested in duplicate). ITC experiments were thus discontinued and SPR (a technique with a significantly reduced protein consumption) was the technique of choice for affinity determination.

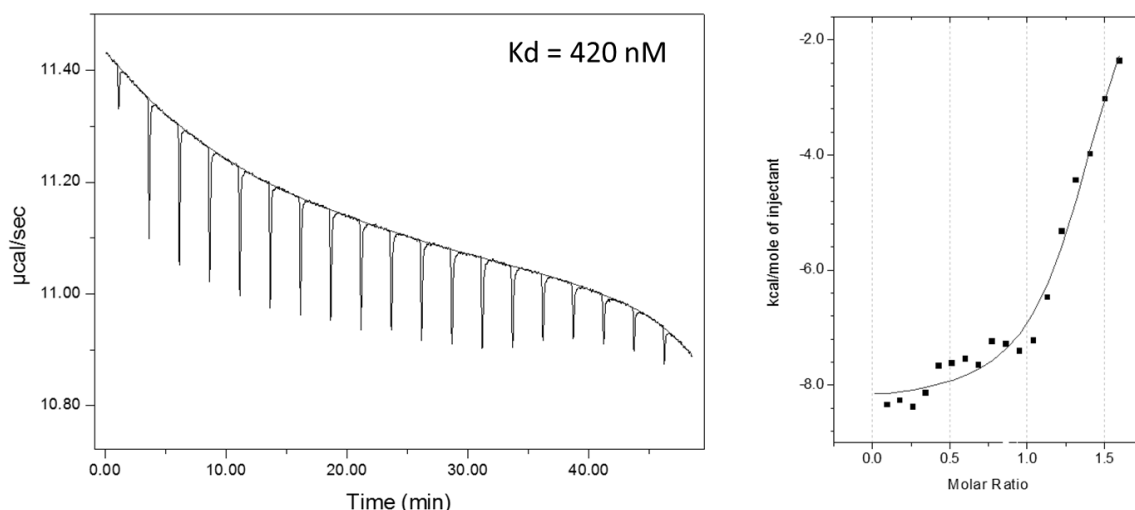


Figure 5.11. Raw (left) and integrated ITC data (right) for the titration of 15 μM SaHPPK with 125 μM **8921**. The K_D value for **8921** binding to ATP/SaHPPK as determined by SPR methods was 330 nM³⁰ and 660 nM.³¹ 1.5 mM AMPCPP was present in both cell and syringe solutions. Difficulties with ITC included the baseline drift observable here in the raw data.

5.11 Differential scanning fluorimetry

The DSF assay assesses the unfolding of a protein as a sample is heated; modifying the chemical components (buffers, ligands etc) present in the protein solution will affect the temperature at which the protein unfolds (the “melt temperature”, T_m). Results from this assay therefore inform conditions that can promote protein stability — this stability, and thus conformational homogeneity, in turn promotes crystallisation.

Fluorescence intensity was measured with excitation/emission = 490/570 nm using a Bio-Rad CFX96 or CFX384 thermocycler. ABgene™ plates (Fisher Scientific) were heated from 25–95 °C with a heating rate of 1.0 °C·min⁻¹. Well volume was 20 μL with 0.3 μL of SYPRO® orange dye and 0.3 μL of protein sample. The Phoenix liquid handling system was used to dispense reagents and protein into plates. Negative controls lacking either protein or dye were present in each plate tested. Additionally, lysozyme was present as a test protein in each plate. Controls and samples were tested in triplicate. Data were acquired on the BioRad CFX Manager (version 3.1) and processed using Meltdown.³² T_m values were calculated from the negative peak of the first derivative of the melt curve.

Compound **SAN3** was initially confirmed to bind to SaHPPK via a DSF assay. This was performed in a solution containing 50 mM HEPES, pH 8.0, 10 mM MgCl₂, and 1 mM TCEP, with protein present at 2 μ M. **SAN3** was diluted from a 100 mM DMSO stock to a concentration of 10–2000 μ M. **SAN3** was also tested over this same concentration range in the presence of 100 μ M AMPCPP.

A DSF screen developed by the C3 centre was used to test the EcDHPS and EcHPPK enzymes (similar tests for SaHPPK being performed previously²⁹). This screen³³ employs a range of buffers at pH 5–9, at both low (50 mM) and high (200 mM) NaCl concentrations. Protein was present at a final concentration of 2 μ M. This assay indicated that, for EcHPPK, the present conditions (20 mM tris-chloride, pH 8.0, 100 mM NaCl) were appropriate. EcDHPS appeared most stable in a solution containing 50 mM tri-sodium citrate, pH 6, contrasting with the previous tris (pH 8.5) buffer. 50 mM tri-sodium citrate, pH 6, was thus used as a base condition in a variant of the Hampton Research Solubility and Stability Screen, prepared by the C3 centre. This additive screen revealed 150 mM NaH₂PO₄ and 350 mM K₂HPO₄ as factors that further improved the stability of EcDHPS. Custom screening around these conditions gave the final conditions of 50 mM tri-sodium citrate, 150 mM NaH₂PO₄, 350 mM K₂HPO₄, pH 6.9.

A ligand binding assay was performed in EcDHPS in the above buffer. Typically 500 μ M of compound was added to give a final DMSO concentration of 2.5%. ΔT_m values were obtained by subtracting the T_m of ligand-bound enzyme to that of apo-enzyme (2.5% DMSO present as control).

5.12 KinaseGlo biochemical assay

The KinaseGloTM assay kit (Promega) was used to quantify the activity of SaHPPK. This assay employs firefly luciferase to produce a luminescence signal proportional to the remaining concentration of ATP after reaction. The consumption of ATP is in this way determined, and thus, the level of enzyme activity. A previously optimised²⁹ SaHPPK concentration of 0.4 ng· μ L⁻¹ assay volume was used. This concentration allows for monitoring of the first 10% of ATP turnover throughout the 20 min assay. Measurements were performed in 96-well plates in an assay buffer containing 100 mM tris-chloride, 10 mM MgCl₂, pH 8.5, 0.01% [w/v] BSA, 0.01% [v/v] Tween 20, and 10 mM β -mercaptoethanol. Typically 5 μ L of test compound (dissolved in 50% DMSO) and 20 μ L of enzyme were added

to each well followed by 25 μ L of assay buffer to give a final concentration of 0.3 μ M HMDP and 0.2 μ M ATP. After 20 minutes the reaction was stopped by addition of 50 μ L of KinaseGlo™ reagent. HMDP is light sensitive and thus dark rooms and aluminium foil were used when possible during both the setup and running of assays. Luminescence was recorded after 10 min with a FLUOstar Optima plate reader (BMG, Labtech Ltd). The positive control consisted of the reaction mixture (without inhibitor) and the negative control consisted of the reaction mixture without HMDP. The luminescence obtained from these controls defined the maximum and minimum reaction, respectively. A control lacking ATP was also present to confirm that the luminescence observed was from ATP. Controls were performed in octuplicate. Reactions containing inhibitor were performed in triplicate. For single-point screens, inhibitors were tested at 100 μ M. When testing the IC₅₀, compounds were typically tested from a top concentration of 200 μ M, with a 2-fold dilution, for a total of 11 points. Kinetic data and inhibition data were then fitted to Michaelis-Menten and sigmoidal dose-response equations respectively, using GraphPad Prism.

5.13 In vitro antibacterial assay.

Minimum inhibitory concentration (MIC) values for inhibitors were determined as per ref 34. A single colony of *E. coli* K-12 strain MG1655 was grown at 37 °C at 250 rpm in M9 minimal salts medium (Appendix I) containing 0.4% glucose and 20 mM ammonium chloride as carbon and nitrogen sources, respectively. The saturated culture was diluted 1/50 into fresh medium and grown at 37 °C to an OD₆₀₀ of ~0.4–0.5. The subculture was diluted 10³-fold into fresh medium and set up to a final volume of 200 μ L in clear flat-bottom 96-well Nunclon™ surface plates (Thermo Fisher Scientific). While growing the above culture, compounds were dispensed into the plate in duplicates with a top (final) concentration of 256 μ M. Upon mixing the bacterial culture with the compounds, the OD₆₀₀ of the plates was read using a SPECTROstar Nano (BMG Labtech). Plates were incubated in the reader for 16 h at 37 °C. After this time, plates were inspected for visible cell culture growth. The OD₆₀₀ measurements aided in distinguishing cell growth from turbidity caused by other factors, e.g. compound insolubility. No compounds displayed inhibition of cell growth in these experiments. Compound testing was subsequently performed in the presence of the trimethoprim, polymyxin B, or polymyxin E (colistin), at pre-determined sub-MIC concentrations. Alongside polymyxin B or E, an antibacterial effect was observed for many of our compounds. In these experiments, the relevant polymyxin (at a concentration of 400

$\mu\text{g}\cdot\text{mL}^{-1}$ in water) was added to the bacterial culture to a final concentration of $0.25\ \mu\text{g}\cdot\text{mL}^{-1}$ prior to dispensing. Compounds that completely inhibited cell growth (at $256\ \mu\text{M}$ in the presence of $0.25\ \mu\text{g}\cdot\text{mL}^{-1}$ polymyxin) were retested in a two-fold dilution series (typically 5 concentrations, i.e. $256\text{--}16\ \mu\text{M}$) to determine the MIC. MIC values for the compounds were typically $128\text{--}256\ \mu\text{M}$ or $>256\ \mu\text{M}$. This assay is exemplified with compound **ML-9-80** in Figure 5.12.

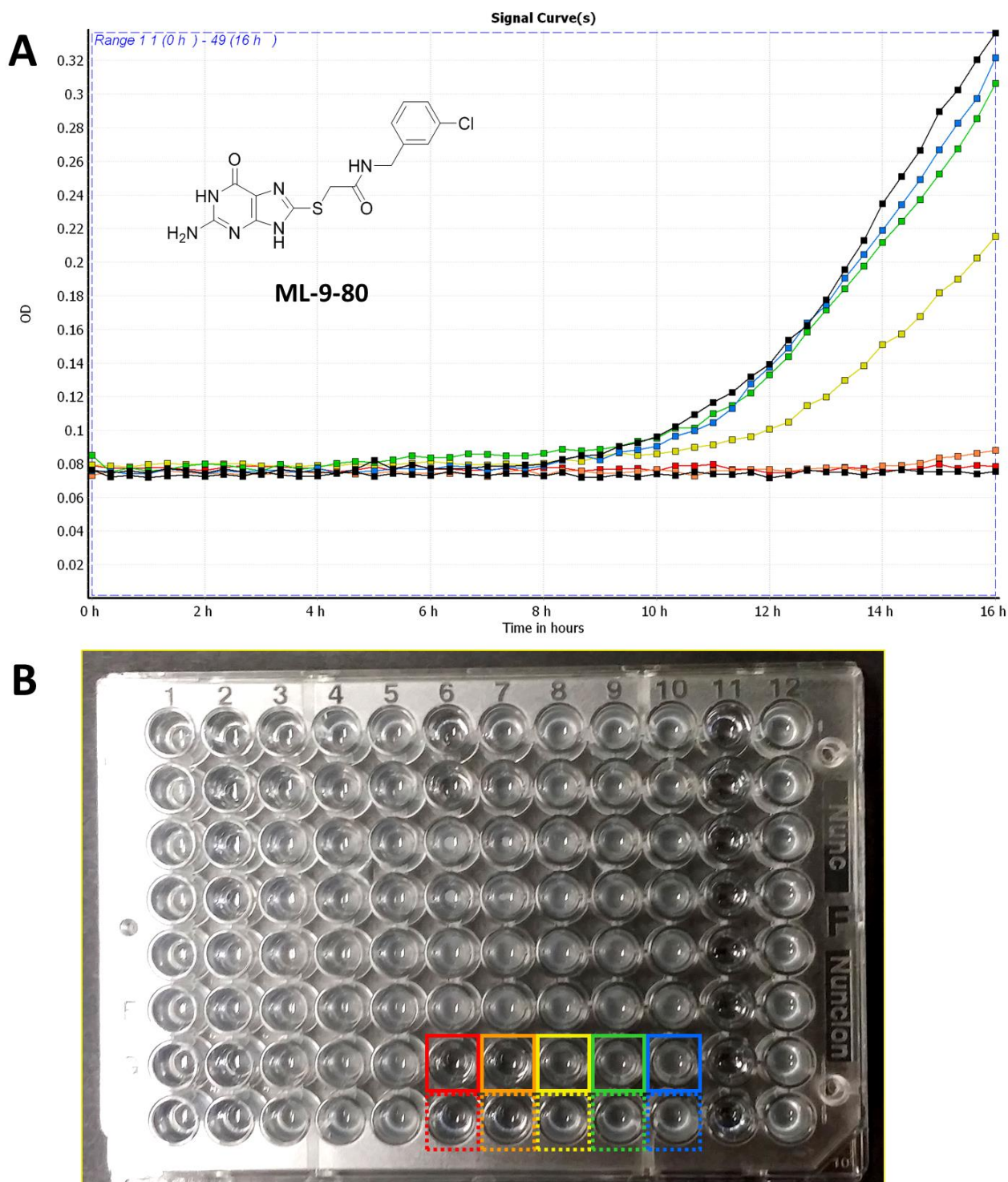


Figure 5.12. Antibacterial activity of compound **ML-9-80** in the presence of $0.25 \mu\text{g}\cdot\text{mL}^{-1}$ colistin. (A) Growth curve measuring OD_{600} over 16 h. Positive and negative controls are displayed in black and **ML-9-80** is displayed at a concentration of $256 \mu\text{M}$ (red), $128 \mu\text{M}$ (orange), $64 \mu\text{M}$ (yellow), $32 \mu\text{M}$ (green), and $16 \mu\text{M}$ (blue). (B) Photo taken after 16 h growth. Colours correspond to those in (A). Columns 11 and 12 contain the sterile control (no cells) and positive growth control (no compound), respectively. Dashed lines indicate the duplicate, which is not displayed in (A) for clarity.

5.14 References

1. Newman, J.; Pham, T.; Peat, T. Phoenix experiments: combining the strengths of commercial crystallization automation. *Acta Crystallogr. Sect F Struct. Biol. Cryst. Commun.* **2008**, *64*, 991–6.
2. Luft, J. R.; DeTitta, G. T. A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55*, 988–993.
3. Cowieson, N. P.; Aragao, D.; Clift, M.; Ericsson, D. J.; Gee, C.; Harrop, S. J.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunnicliffe, A.; Williamson, R.; Caradoc-Davies, T. MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **2015**, *22*, 187–190.
4. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132.
5. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **1994**, *50*, 760–763.
6. Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69*, 1204–1214.
7. Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 432–438.
8. Murshodo; Vagin, A.; Dodson, E. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **1997**, *53*.
9. Adams, P.; Afonine, P.; Bunkóczi, G.; Chen, V.; Davis, I.; Echols, N.; Headd, J.; Hung, L.; Kapral, G.; Grosse-Kunstleve, R.; McCoy, A.; Moriarty, N.; Oeffner, R.; Read, R.; Richardson, D.; Richardson, J.; Terwilliger, T.; Zwart, P. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221.
10. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

11. Joosten, R. P.; Womack, T.; Vriend, G.; Bricogne, G. Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *65*, 176–185.
12. Wlodek, S.; Skillman, A. G.; Nicholls, A. Automated ligand placement and refinement with a combined force field and shape potential. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 741–749.
13. Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B.; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007**, *35*, W375–W383.
14. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980.
15. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
16. Yun, M. K.; Hoagland, D.; Kumar, G.; Waddell, M. B.; Rock, C. O.; Lee, R. E.; White, S. W. The identification, analysis and structure-based development of novel inhibitors of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase. *Bioorg. Med. Chem.* **2014**, *22*, 2157–2165.
17. Sivashanmugam, A.; Murray, V.; Cui, C.; Zhang, Y.; Wang, J.; Li, Q. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* **2009**, *18*, 936–948.
18. Studier, F. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **2005**, *41*, 207–34.
19. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*, Vol. 3. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
20. Kimple, A. J.; Muller, R. E.; Siderovski, D. P.; Willard, F. S. A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance. *Methods in molecular biology (Clifton, N.J.)* **2010**, *627*, 91–100.
21. Chhabra, S.; Barlow, N.; Dolezal, O.; Hattarki, M. K.; Newman, J.; Peat, T. S.; Graham, B.; Swarbrick, J. D. Exploring the chemical space around 8-mercaptoguanine as

a route to new inhibitors of the folate biosynthesis enzyme HPPK. *PLoS One* **2013**, *8*, e59535.

22. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277–293.
23. Bartels, C.; Xia, T. H.; Billeter, M.; Güntert, P.; Wuthrich, K. The program XEASY for computer-supported NMR spectral-analysis of biological macromolecules. *J. Biomol. NMR* **1995**, *6*, 1–10.
24. Goddard, T. D.; Kneller, D. *SPARKY 3 ed.*, San Francisco: University of California, 2001.
25. Maestro, version 10.3, Schrödinger, LLC, New York, NY, 2015.
26. Schrödinger Suite 2015-3 Protein Preparation Wizard; Epik version 3.3, Schrödinger, LLC, New York, NY, 2015; Impact version 6.8, Schrödinger, LLC, New York, NY, 2015; Prime version 4.1, Schrödinger, LLC, New York, NY, 2015.
27. LigPrep, version 3.5, Schrödinger, LLC, New York, NY, 2015.
28. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–49.
29. Chhabra, S.; Dolezal, O.; Collins, B. M.; Newman, J.; Simpson, J. S.; Macreadie, I. G.; Fernley, R.; Peat, T. S.; Swarbrick, J. D. Structure of *S. aureus* HPPK and the discovery of a new substrate site inhibitor. *PLoS One* **2012**, *7*, e29444.
30. Dennis, M. L.; Pitcher, N. P.; Lee, M. D.; DeBono, A. J.; Wang, Z.-C.; Harjani, J. R.; Rahmani, R.; Cleary, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structural basis for the selective binding of inhibitors to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus* and *Escherichia coli*. *J. Med. Chem.* **2016**, *59*, 5248–5263.
31. Dennis, M. L.; Chhabra, S.; Wang, Z.-C.; Debono, A.; Dolezal, O.; Newman, J.; Pitcher, N. P.; Rahmani, R.; Cleary, B.; Barlow, N.; Hattarki, M.; Graham, B.; Peat, T. S.; Baell, J. B.; Swarbrick, J. D. Structure-based design and development of functionalized mercaptoguanine derivatives as inhibitors of the folate biosynthesis pathway enzyme 6-

hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Staphylococcus aureus*. *J. Med. Chem.* **2014**, *57*, 9612–9626.

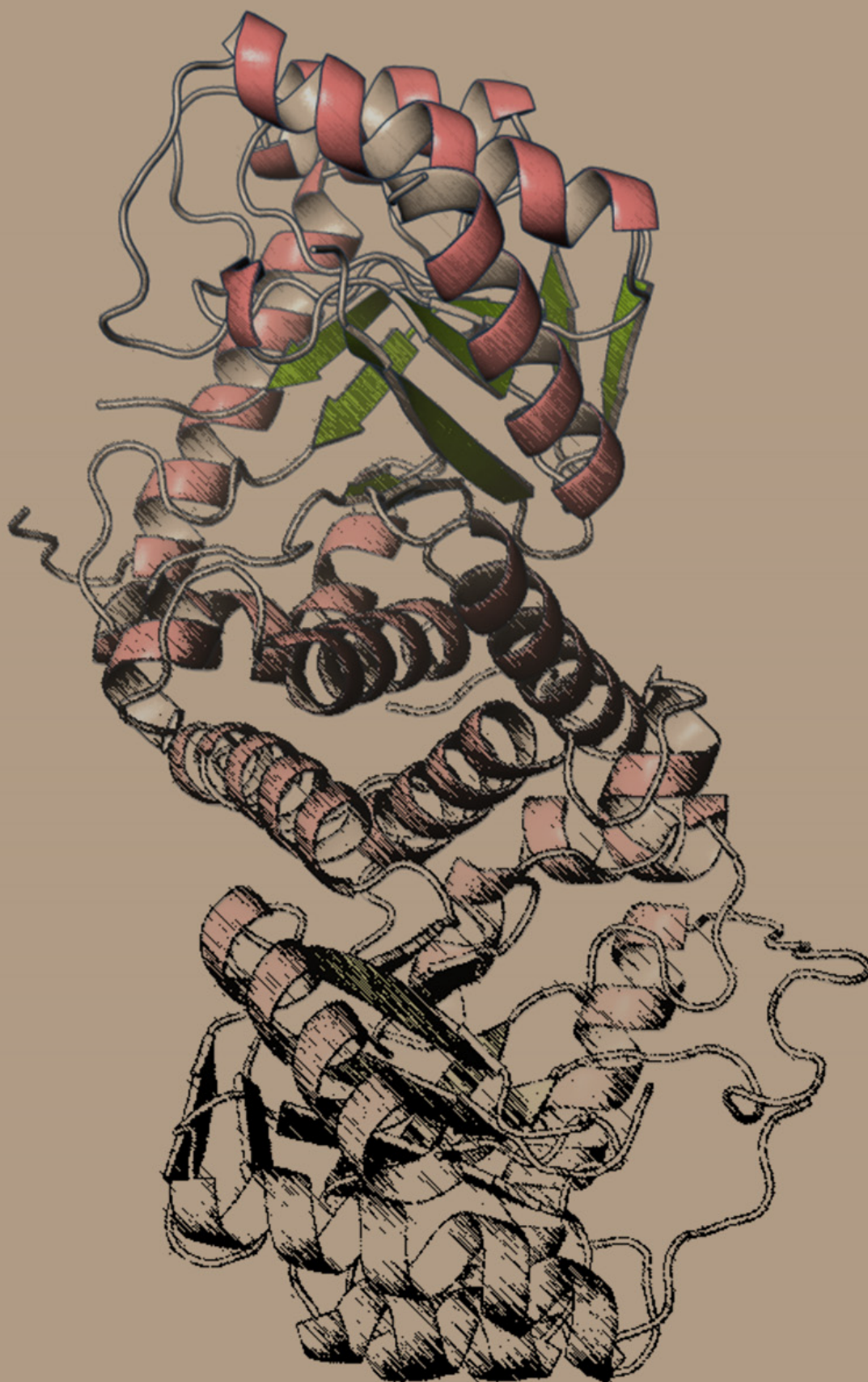
32. Rosa, N.; Ristic, M.; Seabrook, S. A.; Lovell, D.; Lucent, D.; Newman, J. Meltdown: a tool to help in the interpretation of thermal melt curves acquired by differential scanning fluorimetry. *J. Biomol. Screen.* **2015**, *20*, 898–905.

33. Seabrook, S. A.; Newman, J. High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Comb. Sci.* **2013**, *15*, 387–392.

34. Zlitni, S.; Ferruccio, L. F.; Brown, E. D. Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. *Nat. Chem. Biol.* **2013**, *9*, 796–804.

Conclusion

6



This thesis has focused on the development of the fragment 8MG, a folate pathway inhibitor. The derivatives of 8MG presented in this thesis improve upon this fragment's affinity >100-fold for all enzymes investigated (SaHPPK, EcHPPK, and EcDHPS). NMR and X-ray crystallographic studies have thoroughly characterised the binding of these compounds and helped reveal directions that future research can take.

The fortuitous discovery that compound **SAN3** bound to SaHPPK ~10-fold tighter than the parent, 8MG, was the catalyst for much of the work undertaken during this project, spurring an investigation into structure and SAR that involved the design, synthesis, and testing of many compounds. Crystal structures of compounds possessing an 8MG-acetophenone scaffold created an understanding of ligand binding for the tested compounds. Later crystal structures of SaHPPK countermanded some of this understanding, when a novel binding pocket was revealed to form in the presence of 8MG-benzyl ligands. Crystal structures of EcHPPK with the same ligands revealed that formation of this pocket was species-specific to SaHPPK. The previously puzzling SAR and species specificity for SaHPPK was thereby clarified. The compound series comprehensively explored the SAR for this new binding site, with the crystal structures revealing points of attachment for further synthetic elaboration.

Parallel screening of these 8MG derivatives against EcDHPS revealed its own SAR story, with different lead compounds emerging. Crystallisation of protein complexes once again helped develop an understanding of the requirements for binding. Synthesis and testing of a small focused library of compounds yielded a slight improvement in binding, albeit at the cost of ligand efficiency. The current lead compounds possess various points for potential synthetic elaboration, which may lead to more potent EcDHPS inhibitors.

Although lacking antibacterial efficacy when administered alone, a number of the presented compounds were shown to effectively inhibit *E. coli* growth in the presence of polymyxin B or E (colistin). These results validate the pterin site of these folate enzymes as antimicrobial targets. The obtained crystal structures inform future synthetic directions that may be pursued to improve potency, however the antimicrobial assay results indicate that research must include a focus on improving cell permeability. Modifying the highly polar 8MG core is likely a key starting point in this regard.

Protein crystallisation represented one of the key challenges throughout this thesis. Fortunately, following optimisation, both EcHPPK and EcDHPS were able to be crystallised with relative consistency. The SaHPPK enzyme, however, continues to represent a

challenge for every new ligand to be co-crystallised. All of the solved SaHPPK complex structures possess key crystal contacts in catalytic loop regions, with conformations dependent on the ligand in question; crystal formation is thus highly ligand-selective. Future research may therefore benefit from modifying the SaHPPK construct. Rational mutation of surface residues to reduce entropy is one method that promotes stable crystal contacts.¹ Mutations made in this way may improve the robustness of crystal growth and diffraction, and facilitate research into HPPK inhibitors.

In summary, the work presented has advanced the field of HPPK and DHPS inhibitor design and revealed avenues for further development of these compounds. A single reaction step was required for synthesis of the majority of compounds presented. This greatly facilitated larger and more comprehensive compound library synthesis and was critical to the success achieved. This work therefore attests to the importance of judicious fragment selection at the start of a drug discovery project in regards to synthetic amenability. The current lead inhibitors remain attractive in this regard, and have potential to be further developed into efficacious antibacterial agents, to join the arsenal of drugs needed to combat antibiotic resistance.

1. Cooper, D. R.; Boczek, T.; Grelewska, K.; Pinkowska, M.; Sikorska, M.; Zawadzki, M.; Derewenda, Z. Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr. D Biol. Crystallogr.* **2007**, 65, 636–645.

Appendix I – Media recipes

SDS-PAGE sample buffer — 400 μ L water (400 μ L β -mercaptoethanol for reduced sample buffer), 1 mL 0.5 M Tris-HCl (pH 6.8), 800 μ L glycerol, 1.6 mL 10% [w/v] SDS, 250 μ L 0.5% [w/v] bromophenol blue

Media (per litre):

LB, rich medium — 10 g tryptone, 5 g yeast extract, 5 g NaCl

2x YT, rich medium — 16 g tryptone, 10 g yeast extract, 5 g NaCl

ZYM-5052, general purpose autoinduction medium — 1 mL 1 M MgSO_4 , 20 mL 50x 5052, 50 mL 20x NPS, 1 mL 50 $\text{mg}\cdot\text{mL}^{-1}$ kanamycin, LB to 1 L.

N-5052, minimal autoinduction media suitable for ^{15}N -labelling — 1 mL 1 M MgSO_4 , 50x 5052, 50 mL 20x ^{15}NPS , 1 mL 10 $\text{mg}\cdot\text{mL}^{-1}$ thiamine, 1 mL 10 $\text{mg}\cdot\text{mL}^{-1}$ biotin, 1 mL trace metal mix, 1 mL 50 $\text{mg}\cdot\text{mL}^{-1}$ kanamycin, MilliQ H_2O to 1 L.

Non-expressing minimal media — 1 mL MgSO_4 , 0.5 % [w/v] glucose (diluted from a 25% [w/v] stock), 50 mL 20x NPS, 1 mL 10 $\text{mg}\cdot\text{mL}^{-1}$ thiamine, 1 mL 10 $\text{mg}\cdot\text{mL}^{-1}$ biotin, 1 mL trace metal mix, 1 mL 50 $\text{mg}\cdot\text{mL}^{-1}$ kanamycin, MilliQ H_2O to 1 L.

M9 minimal media, suitable for ^{15}N - and/or ^{13}C -labelling — 100 mL 10x M9 salts, 1.5 g $^{15}\text{NH}_4\text{Cl}$, 3 g of ^{13}C -glucose, 1 mL 0.1 M CaCl_2 , 1 mL 1 M MgSO_4 , 100 1 mL 10 $\text{mg}\cdot\text{mL}^{-1}$ thiamine, 1 mL 10 $\text{mg}\cdot\text{mL}^{-1}$ biotin, 1 mL trace metal mix, 1 mL 50 $\text{mg}\cdot\text{mL}^{-1}$ kanamycin, MilliQ H_2O to 1 L.

Ingredients:

20xNPS (per 100mL) — 90 mL MilliQ H_2O , 6.6g $(\text{NH}_4)_2\text{SO}_4$, 13.6 g KH_2PO_4 , 14.2 g Na_2HPO_4

20x ^{15}NPS (per 100mL) — 5 g $^{15}\text{NH}_4\text{Cl}$, 1.42 g Na_2SO_4 , 13.6 g KH_2PO_4 , 14.2 g Na_2HPO_4

50x5052 (per 100mL) — 25 g glycerol, 73 mL MilliQ H_2O , 2.5 g glucose, 10 g alpha-lactose.

Trace metal mix (per 100 mL) — all stocks are prepared in MilliQ H₂O except for FeCl₃, which is dissolved in 0.1 M HCl (final HCl concentration is 50 mM). 1.3515 g FeCl₃·6H₂O (dissolved in 50 mL 0.1M HCl), 221.98 mg CaCl₂, 197.91mg MnCl₂·4H₂O, 287.56 mg ZnSo₄·7H₂O, 47.59 mg CoCl₂·6H₂O, 34.098 mg CuCl₂·2H₂O, 47.544 mg NiCl₂·6H₂O, 52.606 mg Na₂MnO₄, 12/366 mg Na₂SeO₃.

M9 salts (per litre) — 67.8 g Na₂HPO₄, 30 g KH₂PO₄, 5g NaCl, adjusted to pH 7.4.

Appendix II – Additional publications during the thesis period

The following papers contain additional work performed throughout the thesis period, namely the expression and purification of both wildtype and mutant forms of ubiquitin and SaHPPK, for use in lanthanide-binding tag studies. The three papers presented are:

1. Lee, M. D.; Loh, C. T.; Shin, J.; Chhabra, S.; **Dennis, M. L.**; Otting, G.; Swarbrick, J. D.; Graham, B. Compact, hydrophilic, lanthanide-binding tags for paramagnetic NMR spectroscopy. *Chemical Science* **2015**, 6, 2614-2624.
2. Swarbrick, J. D.; Ung, P.; **Dennis, M. L.**; Lee, M. D.; Chhabra, S.; Graham, B. Installation of a Rigid EDTA-Like Motif into a Protein α -Helix for Paramagnetic NMR Spectroscopy with Cobalt(II) Ions. *Chemistry – A European Journal* **2016**, 22, 1228-1232.
3. Lee, M. D.; **Dennis, M. L.**; Swarbrick, J. D.; Graham, B. Enantiomeric two-armed lanthanide-binding tags for complementary effects in paramagnetic NMR spectroscopy. *Chemical Communications* **2016**, 52, 7954-7957.

Only the main body of each publication is reprinted. If the supplementary information is desired, this is freely available online and readers are advised to refer to each paper for details on how this can be accessed.



Cite this: *Chem. Sci.*, 2015, 6, 2614

Compact, hydrophilic, lanthanide-binding tags for paramagnetic NMR spectroscopy†

M. D. Lee,^a C.-T. Loh,^b J. Shin,^a S. Chhabra,^a M. L. Dennis,^a G. Otting,^b J. D. Swarbrick^{*a} and B. Graham^{*a}

The design, synthesis and evaluation of four novel lanthanide-binding tags for paramagnetic NMR spectroscopy are reported. Each tag is based on the ((2*S*,2'*S*,2''*S*,2'''*S*)-1,1',1'',1'''-(1,4,7,10-tetraazacyclododecane-1,4,7,10-tetrayl)tetrakis(propan-2-ol)) scaffold, featuring small chiral alcohol coordinating pendants to minimise the size and hydrophobic character of each tag. The tags feature different linkers of variable length for conjugation to protein via a single cysteine residue. Each tag's ability to induce pseudocontact shifts (PCS) was assessed on a ubiquitin A28C mutant. Two enantiomeric tags of particular note, **C7** and **C8**, produced significantly larger $\Delta\chi$ -tensors compared to a previously developed tag, **C1**, attributed to the extremely short linker utilised, limiting the mobility of the bound lanthanide ion. The **C7** and **C8** tags' capacity to induce PCSs was further demonstrated on GB1 Q32C and 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) S112C/C80A mutants. Whilst factors such as the choice of lanthanide ion, pH and site of conjugation influence the size of the PCSs obtained, the tags represent a significant advance in the field.

Received 16th December 2014
Accepted 25th February 2015

DOI: 10.1039/c4sc03892d

www.rsc.org/chemicalscience

Introduction

The site-specific incorporation of paramagnetic metal ions into proteins allows access to unique NMR parameters that can provide valuable structural insights into protein structure and dynamics.^{1–3} These include pseudocontact shifts (PCS), residual dipolar couplings (RDC) and paramagnetic relaxation enhancement (PRE). PCSs are particularly attractive structural restraints as they are simple to measure (as the difference in chemical shift between a diamagnetic and paramagnetic sample) and encompass both distance and orientation information of nuclei relative to the magnetic susceptibility

anisotropy ($\Delta\chi$) tensor. The PCS of any nuclear spin can be back-calculated from knowledge of the $\Delta\chi$ -tensor:

$$\delta_{\text{PCS}} = \frac{1}{12\pi r^3} \left[\Delta\chi_{\text{ax}} (3 \cos^2 \theta - 1) + \frac{3}{2} \Delta\chi_{\text{rh}} \sin^2 \theta \cos 2\varphi \right]$$

where $\Delta\chi_{\text{ax}}$ and $\Delta\chi_{\text{rh}}$ are the axial and rhombic components of the $\Delta\chi$ -tensor and r , θ and φ are the polar coordinates of the nuclei with respect to the principal axes of the $\Delta\chi$ -tensor. The r^{-3} distance dependence of PCSs allows them to be measured for nuclei up to 40 Å or more away from the metal ion.⁴ Thus, PCSs provide long-range structural information that can be utilised in the study of protein structure and conformation,^{5–9} protein–protein^{10–13} and protein–small molecule interactions,^{14–17} and even *de novo* protein structure determination.^{18,19}

Paramagnetic lanthanide(III) ions, except Gd^{3+} , can be used to induce PCSs in the NMR spectra of macromolecules. Their anisotropic magnetic susceptibilities are inherently large (yet different) and, combined with their similar structure and bonding, allow the substitution of one lanthanide ion for another as a convenient route to vary the magnetic properties of a sample. However, most proteins do not natively bind lanthanide ions, which has spurred recent interest in the design of synthetic lanthanide-binding tags (LBTs)^{3,20} or peptides^{21,22} capable of introducing lanthanide ions into proteins in a site-specific manner.

Lanthanide ions are “hard” Lewis acids that can adopt high coordination numbers, thus polydentate ligands featuring hard bases (such as O and N atoms) are ideal candidates to ensure tight lanthanide ion binding. Rigidity of the lanthanide relative to

^aMonash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia

^bResearch School of Chemistry, Australian National University, Canberra, ACT 0200, Australia

† Electronic supplementary information (ESI) available: Syntheses of **C5** and **C6**, ¹H NMR spectra and HRMS of Yb^{3+} complexes of **C5**–**C7**, experimental details for expression, purification and tagging of GB1 Q32C and HPPK S112C/C80A, ¹⁵N-HSQC spectra of differently tagged ubiquitin A28C, GB1 Q32C and HPPK S112C/C80A, correlation between PCSs of **C7** tagged ubiquitin A28C at pH 8 and pH 6.5, comparison of measured PCSs against residue number for **C5**–**C8** tagged ubiquitin A28C, isosurface representations of $\Delta\chi$ -tensors of **C1**, **C7** and **C8** tagged ubiquitin A28C, simultaneously calculated $\Delta\chi$ -tensor properties for tagged ubiquitin A28C and GB1 Q32C, experimental RDCs and alignment tensor properties for **C7**-**Tm**³⁺ and **C8**-**Tm**³⁺ tagged ubiquitin A28C, experimental PCSs of tagged ubiquitin A28C, GB1 Q32C and HPPK S112C/C80A, 1D ¹H and ¹³C NMR spectra of novel small molecule compounds, analytical HPLC traces of **C5**–**C7**. See DOI: 10.1039/c4sc03892d



the protein frame is paramount to prevent the deleterious averaging effects of tag mobility on measured PCSs and RDCs, which tend towards zero with increasing motion. Engineering tagging sites to take advantage of additional coordination to acidic side-chains of proteins,^{23–25} conjugation to proteins through multiple sites of attachment^{26,27} or the use of steric bulk²⁸ have been successful strategies to limit tag mobility. It is imperative that the attached LBT must also give rise to a single observable species in solution, as multiple species in *slow* exchange can lead to highly complex spectra that are of limited practical utility.^{29–31}

LBTs that bind lanthanide ions extremely tightly, without the need for additional protein interactions, are particularly attractive. They allow the study of proteins in the presence of their own native metal ions and metal ion-bound cofactors (*e.g.* metalloproteins and kinases) and remove any problems associated with excess free lanthanide ions that can result in line broadening in the NMR spectra. LBTs based on DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) have proved to be useful, having induced significant PCSs and RDCs in several proteins.^{28,32,33} Although capable of binding lanthanides with dissociation constants of the order of 10^{-23} to 10^{-25} M,³⁴ lanthanide complexes of DOTA display a dynamic behaviour in solution at ambient temperature. Inversion of the cyclen ring (defined by the NCCN torsion angle as either $\delta\delta\delta\delta$ or $\lambda\lambda\lambda\lambda$) and rotation of the pendant arms (defined by the NCCO torsion angles as either Δ or Λ) result in a dynamic equilibrium between square anti-prismatic (SAP) and twisted square anti-prismatic (TSAP) coordination geometries.³⁵ When bound to a protein, this can lead to the presence of multiple stereoisomers in *slow* exchange, each producing their own paramagnetic effects that greatly complicate analysis of the spectra. In order to limit these conformational exchange processes and simplify the spectra, successful DOTA-based LBT designs have incorporated chiral elements into the pendant arms or cyclen ring,^{28,33} or employed two-points of conjugation to the protein.^{32,36}

Previously, we demonstrated that attachment of three sterically bulky (*S*)- or (*R*)-phenethylacetamide pendant arms to a 1,4,7,10-tetraazacyclododecane (cyclen) macrocyclic ring (C1–C4, Fig. 1) was sufficient to generate a single apparent stereoisomer and to limit tag flexibility, allowing the observation of measurable and sizable paramagnetic effects.^{28,37} However, for some proteins in our laboratory, such as 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), this family of tags was found to present issues in terms of protein stability, as evidenced by an increased tendency to precipitate during and post conjugation. We suggest this to originate from the incompatibility of the large hydrophobic nature of this tag series with these proteins. It is also foreseeable that the tags' hydrophobic character could complicate the study and screening of weak ligand–protein interactions, as small hydrophobic compounds (*e.g.* from fragment libraries) can potentially associate transiently with the LBT, leading to a transferred PCS effect and a skewed, “meaningless” average ligand PCS.

With this in mind, we have now developed a new series of tags that are much more structurally compact and hydrophilic in nature (C5–C8, Fig. 1). These tags are based on ((2*S*,2'*S*,2''*S*,2'''*S*)-1,1',1'',1'''-(1,4,7,10-tetraazacyclododecane-

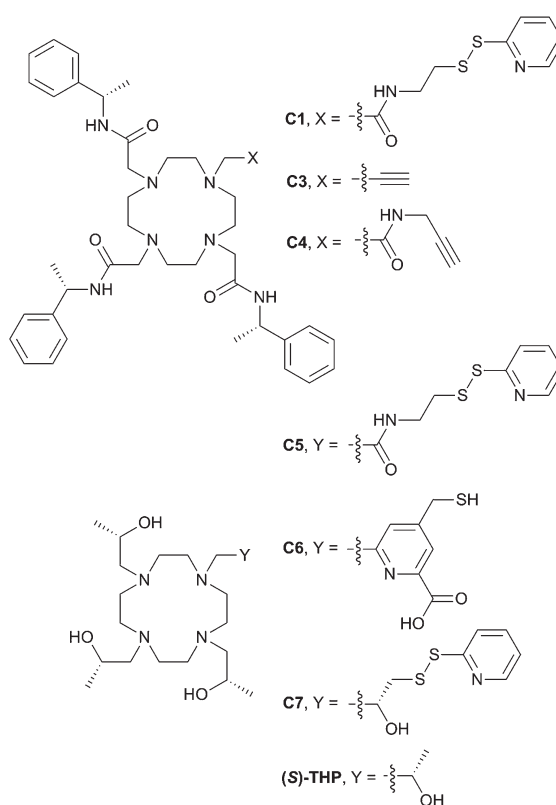


Fig. 1 Existing and newly developed LBTs referred to in the text. C2 and C8 are the enantiomers of C1 and C7, respectively.

1,4,7,10-tetrayl)tetrakis(propan-2-ol)) ((*S*)-THP), a cyclen derivative featuring four chiral (*S*)-2-hydroxypropyl pendants. Multiple (*S*)-THP- Ln^{3+} (where Ln = La, Ce, Nd, Eu, Yb or Lu) complexes have been reported to show ^1H NMR spectra that display a single set of resonances,^{38–40} which suggested that an (*S*)-THP based LBT could also produce a single set of PCSs to nuclei of a bound protein. The (*S*)-THP- Yb^{3+} complex specifically has been shown to adopt a $\Lambda(\lambda\lambda\lambda)$ TSAP geometry in solution.³⁹

Conjugation of single-point attachment LBTs to proteins requires less prior structural knowledge of the target, fewer mutations for their introduction and can still produce useful effective $\Delta\chi$ -tensors when tag movements are limited.⁴¹ Thus, our initial focus has been on the development of (*S*)-THP derivatives featuring a single thiol-conjugatable group, so as to produce tags applicable to the study of as wide a range of protein systems as possible. The first of these (C5) utilises the same pyridyl disulfide-activated linker as our earlier reported C1 and C2 tags. Given the absence of the sterically bulky pendants of the latter tags, which were postulated to be an important element in limiting tag flexibility,²⁸ it was anticipated that this tag might prove too mobile for NMR applications. Therefore, analogues with shorter linker groups were also engineered. C6 features a bidentate chelating 2-carboxypyridine moiety with a conjugatable methylmercaptan group attached to the 4-position of the pyridine ring, and can be viewed as a hybrid of (*S*)-THP and the various DPA-based LBTs reported by Otting and co-





workers.^{23,42–44} C7, and its enantiomer C8, feature possibly the smallest practicable modification to (S)-THP that allows for bioconjugation: a pyridyl disulfide group is attached directly to one of the four chiral 2-hydroxypropyl pendants and the resulting protein-conjugated tags feature only a disulfide bond between the chirally pure (S)/(R)-THP-Ln³⁺ chelate and protein.

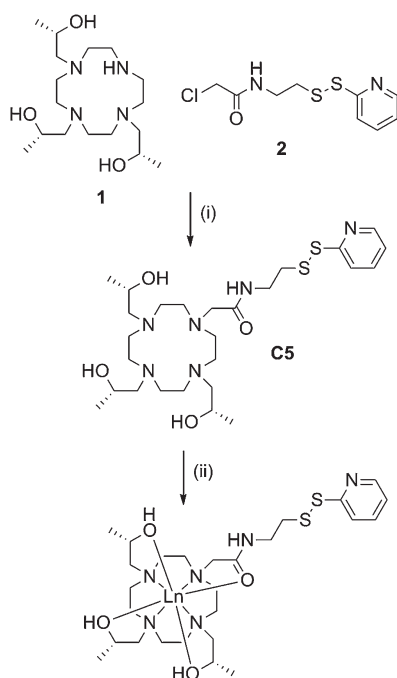
We now report the synthesis of the new tags (C5–C8) and demonstrate their utility in paramagnetic NMR structural studies using human ubiquitin and GB1 as model proteins, as well as the antimicrobial target, HPPK.⁴⁵ As detailed below, the C5 and C6 tags are found to perform comparably to C1 in terms of the magnitude of the $\Delta\chi$ -tensors observed on ubiquitin. More significantly, however, the C7 and C8 tags produce considerably larger paramagnetic effects, indicating that the short linker present within these tags translates to a more restricted lanthanide ion attachment to the protein.

Results and discussion

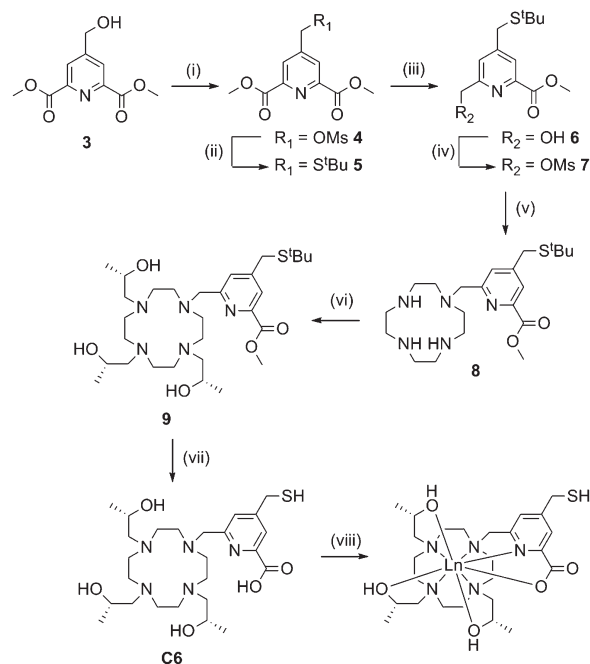
Tag synthesis

C5 was prepared in good overall yield by nucleophilic substitution between the previously reported compounds, (1S,4S,7S)-1,4,7-tris(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane⁴⁶ (**1**) and 2-chloro-N-(2-(pyridin-2-ylthio)ethyl)acetamide.²⁸ (**2**) (Scheme 1).

Synthesis of C6 (Scheme 2) required preparation of a novel carboxyl pyridine linker. Dimethyl 4-(hydroxymethyl)pyridine-2,6-dicarboxylate (**3**) was prepared following literature procedures²³ and converted to the *tert*-butyl thioether **5** via the



Scheme 1 Synthesis of C5 and its lanthanide complexes. Reagents and conditions: (i) DIPEA, ACN, RT, 72 h, 61%; (ii) LnCl₃, ACN, H₂O, pH 7, reflux, overnight, quant.



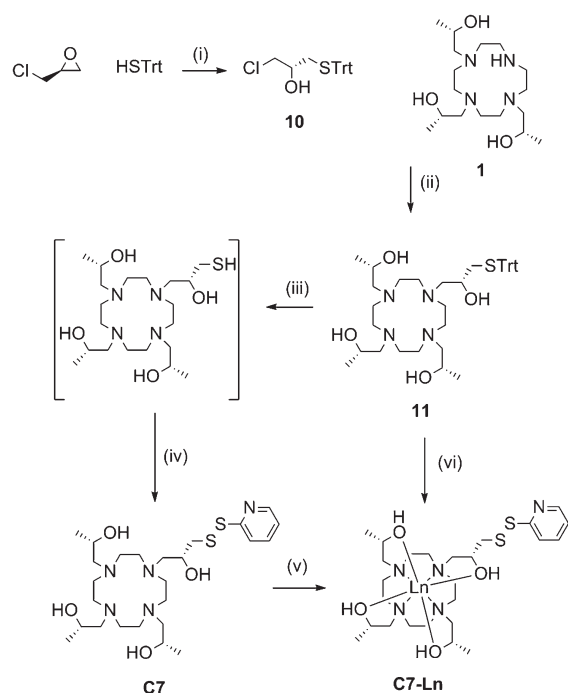
Scheme 2 Synthesis of C6 and its lanthanide complexes. Reagents and conditions: (i) MsCl, DIPEA, DCM, 0 °C, 30 min, quant.; (ii) ^tBuSH, NaH, DMF, RT, 5 min, 47%; (iii) NaBH₄, MeOH, DCM, RT, 2 h, 64%; (iv) MsCl, DIPEA, DCM, 0 °C, 30 min, 77%; (v) cyclen, CHCl₃, RT, O/N, quant.; (vi) (S)-propylene oxide, MeOH, RT, 48 h, quant.; (vii) HCl (32%), reflux, 4 h, 85%; (viii) LnCl₃, ACN, H₂O, pH 7, reflux, overnight, quant.

mesylate derivative **4**. Partial reduction with sodium borohydride and mesylation of the resulting hydroxyl group yielded **7**, which was reacted with an excess of cyclen to form **8**. Reaction with an excess of (S)-propylene oxide, followed by ester and *tert*-butyl deprotection yielded **C6**. We attempted to activate the thiol of **C6** as a pyridyl disulfide, however the resulting product was unstable during purification, thus the free thiol was used for tagging (*vide infra*).

Metal complexes of C5 and C6 were prepared by heating the relevant tag with two equivalents of XCl₃ salts (X = Y, Dy, Tb, Tm or Yb) at 80 °C in a water–acetonitrile mixture buffered at neutral pH overnight. Coordination of these tags was generally close to quantitative, with excess metal ions and uncomplexed tag removed *via* HPLC purification.

Due to the favourable properties of C7 (*vide infra*), its synthesis underwent several iterations in order to improve the yield (Scheme 3). Similarly to C5 and C6, the initial method involved synthesis of the tag, followed by metal ion complexation.

Ring opening of (S)-epichlorohydrin with triphenylmethane thiol, in the presence of potassium fluoride, produced **10** in excellent yield (95%). **10** then underwent nucleophilic substitution by heating with **1** and potassium carbonate to form **11**. Deprotection of the trityl group was carried out at room temperature with trifluoroacetic acid and triethylsilane. Subsequent thiol activation with 2,2'-dipyridyl disulfide and purification *via* HPLC produced C7 in 39% yield from **11**.



Scheme 3 Synthesis of **C7** and its lanthanide complexes. Reagents and conditions: (i) KF, MeOH, RT, 72 h, 95%; (ii) K_2CO_3 , ACN, reflux, overnight, 60%; (iii) TFA, triethylsilane, DCM, RT, 1 h; (iv) 2,2'-dipyridyl disulfide, MeOH, RT, 15 min, 39% (from **11**); (v) $LnCl_3$, EtOH, DIPEA, reflux, overnight; (vi) $LnCl_3$, MeOH, reflux, 4 h, 2,2'-dipyridyl disulfide, silver nitrate, RT, 2 h, 34%.

Formation of **C7-Ln**³⁺ complexes was extremely slow in the presence of water and required heating in anhydrous ethanol. Complexation was still relatively slow compared to the other tags. Furthermore, if left for a prolonged period of time (*e.g.* greater than 48 h) noticeable amounts of disulfide rearrangement would occur, resulting in a chelate dimer and regeneration of 2,2'-dipyridyl disulfide. **C7-Ln**³⁺ complexes formed this way were thus generally purified from a mixture with uncomplexed **C7**, before a significant amount of disulfide rearrangement could occur, resulting in relatively poor yields.

Various attempts to optimise **C7** complexation were made, including initial passage of **C7** over anion exchange resin (to remove trifluoroacetic acid, present from prior HPLC purification) and addition of organic or inorganic bases to complexation reactions. However, we eventually found the most practical way of producing **C7-Ln**³⁺ complexes to be by forming metal complexes of **11**, before trityl deprotection and thiol activation to the final product. Compound **11** was isolated as a neutral compound and readily formed **11-Ln**³⁺ without side-product formation, by heating for several hours in methanol with two equivalents of the relevant metal chloride salt. **11-Ln**³⁺ was then trityl deprotected with silver nitrate and thiol activated with 2,2'-dipyridyl disulfide, before reverse-phase HPLC purification to yield **C7-Ln**³⁺. This method allowed the formation of **C7-Ln**³⁺ complexes from **11** in "one pot" and required one less round of HPLC purification compared to the previous route, resulting in

overall higher yields (34% from **11**). The **C8** tag and complexes followed the same procedures with the replacement of (*S*)-propylene oxide and (*S*)-epichlorohydrin with their (*R*)-enantiomers.

Fig. S1–S3† show the ¹H NMR spectra of the Yb³⁺ complexes of **C5**–**C7**. Although greatly complicated by the pyridyl disulfide linker breaking the symmetry of the complex, the ¹H NMR spectrum of **C7-Yb**³⁺ bears some resemblances to that of the (*S*)-THP-Yb³⁺ complex reported by Lelli *et al.*³⁹ Comparing the most resolved signals, the peak at –28 ppm in (*S*)-THP-Yb³⁺ is split into four overlapping peaks of equal intensity in **C7-Yb**³⁺, while the peak at 52 ppm in (*S*)-THP-Yb³⁺ is split into three peaks in **C7-Yb**³⁺, one of which is twice the intensity of the other two. The ¹H NMR spectra of the more structurally-varied **C5-Yb**³⁺ and **C6-Yb**³⁺ complexes show fewer similarities to (*S*)-THP-Yb³⁺. We did not attempt a complete assignment of the ¹H NMR spectra of the **C5**–**C7** complexes.

Testing of tags on a cysteine-bearing mutant of ubiquitin

A human ubiquitin A28C mutant was used as an initial model protein to assess the effects of the paramagnetic properties of each tag. Purified protein was first stirred with ten equivalents of DTT to reduce any oxidised cysteines. Excess DTT was removed by passage over a PD10 column equilibrated with 50 mM HEPES, pH 8.0. For the pyridyl disulfide-containing tags, **C5**, **C7** and **C8**, five equivalents of the relevant lanthanide-complexed tag were added and the solutions stirred for 2 h at room temperature, before excess tag was removed by passage over a PD10 column. Tagging yields varied between 70% to quantitative, as determined by NMR analysis.

In order to conjugate **C6**, reduced protein was first reacted with a ten-fold excess of DTNB for 1 h, before passage over a PD10 column followed by the addition of five equivalents of **C6-Ln**³⁺ complex. The reaction was allowed to stir for 2 h at room temperature, before removal of excess tag *via* a PD10 column. Tagging yields were generally quantitative.

¹⁵N-HSQC spectra of each lanthanide complex conjugated to ubiquitin A28C showed significant PCSs (Fig. 2 and S7–S9†). For each tag, the Y³⁺ complex produced minor chemical shift perturbations relative to the untagged protein, with larger shifts limited to residues in the vicinity of the tagging site. In each spectrum, only a single set of PCSs was observed. PCSs were measured as the difference in chemical shift of resonances between the paramagnetic (Dy³⁺, Tb³⁺, Tm³⁺ or Yb³⁺) and diamagnetic (Y³⁺) tagged samples. The $\Delta\chi$ -tensors were determined by fitting the measured PCSs (Tables S2 and S3†) to the first conformer of the NMR structure of ubiquitin (PDB ID 2MJB),⁴⁷ both individually for each metal ion (Table 1) and simultaneously for each complex of a given tag with a common metal ion position (Table S4,† *vide infra*). Fig. 3 shows the correlations between measured and back-calculated PCSs for the individually derived $\Delta\chi$ -tensors, demonstrating their high quality, which is also reflected in the low *Q*-values.

Different *pK_a* values for the deprotonation of a single alcohol pendant (or, possibly, bound water molecule) have been reported for different (*S*)-THP-Ln³⁺ complexes. These range



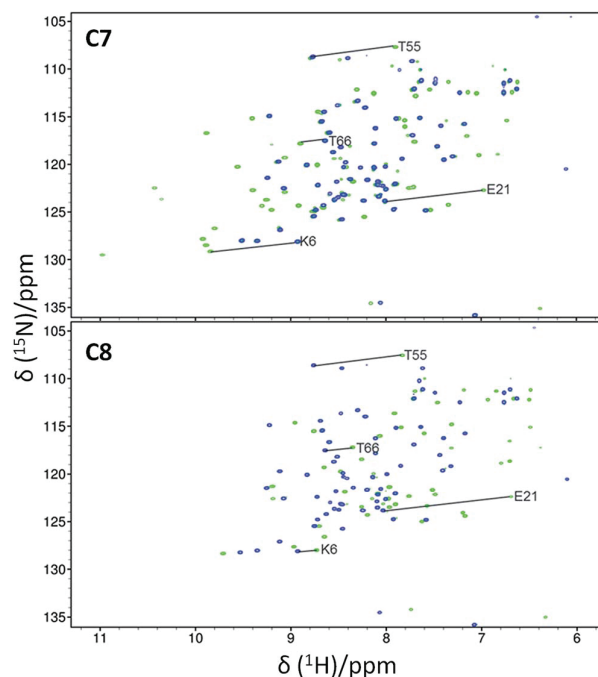


Fig. 2 Overlays of ^{15}N -HSQC spectra of C7 (top spectra) and C8 (bottom spectra) tagged ubiquitin A28C, loaded with either Y^{3+} (blue) or Tm^{3+} (green). The spectra were recorded at 25 °C and pH 8.0 at a ^1H NMR frequency of 600 MHz. Selected PCSs are indicated with solid lines.

from 8.4 for the lanthanum complex to 6.4 for lutetium, with a trend of decreasing pK_a across the lanthanide series.^{39,48} Thus, at a given pH, different (S)-THP- Ln^{3+} complexes can exist across a range of equilibria between +3 and +2 charged states.

Table 1 $\Delta\chi$ -Tensor parameters for C5–C8 tagged ubiquitin A28C^{a,b}

Tag	Ln^{3+}	# PCS	$\Delta\chi_{\text{ax}}$	$\Delta\chi_{\text{rh}}$	Q	x	y	z	α	β	γ
C5	Dy^{3+}	39	8.2	5.3	0.04	2.908	2.285	−15.138	141	88	71
	Tb^{3+}	47	9.4 (0.5)	2.2 (0.1)	0.06	2.308	−0.421	−17.179	157	95	114
	Tm^{3+}	47	−18.7 (1.7)	−6.9 (0.4)	0.06	4.728	−3.051	−17.815	127	96	100
	Yb^{3+}	61	−6.7 (0.4)	−2.1 (0.2)	0.08	0.857	−2.115	−18.247	120	97	122
C6	Dy^{3+}	49	−9.4	−5.9	0.07	8.694	3.797	−11.227	45	49	80
	Tb^{3+}	47	−14.6 (0.4)	−3.6 (0.1)	0.04	7.019	2.304	−13.622	44	68	97
	Tm^{3+}	51	11.5 (0.7)	4.1 (0.4)	0.10	7.018	3.097	−12.773	41	71	130
	Yb^{3+}	51	2.0 (0.1)	0.9 (0.0)	0.08	9.178	2.136	−12.736	42	60	103
C7	Dy^{3+}	35	26.6 (1.1)	6.0 (0.5)	0.03	−0.734	−3.238	−13.305	71	29	32
	Tb^{3+}	40	11.7 (0.3)	1.7 (0.3)	0.04	−1.901	−3.712	−14.371	47	44	59
	Tm^{3+}	44	−19.4 (0.7)	−7.8 (0.9)	0.03	−4.314	−1.357	−13.717	9	62	104
	Yb^{3+}	51	5.8	3.0	0.03	−0.427	−0.482	−14.129	16	110	85
C8	Dy^{3+}	28	31.2 (0.7)	7.1 (0.6)	0.02	1.810	−3.922	−13.760	91	44	5
	Tb^{3+}	37	14.3 (0.6)	5.1 (0.7)	0.05	2.003	−1.928	−13.959	73	38	19
	Tm^{3+}	43	−16.3	−10.2	0.04	0.862	−3.716	−15.075	67	26	46
	Yb^{3+}	46	−4.2 (0.1)	−1.9 (0.2)	0.04	1.679	−3.184	−14.267	140	22	169

^a The axial and rhombic components of the $\Delta\chi$ -tensors are reported in units of 10^{-32} m^3 , and the Euler angles in degrees, using the yz convention and unique tensor representation.⁵³ Standard deviations (in brackets) were determined from random removal of 10% of the PCSs and recalculating the $\Delta\chi$ -tensor 1000 times, in some cases the z and y axes of the tensor were of similar magnitude and swapped in different fits, thus standard deviations were not determined. Quality factors (Q) were calculated as the root-mean-square deviation between the experimental and back-calculated PCSs divided by the root-mean-square of the experimental PCSs. ^b Metal ion coordinates (x, y, z) are reported relative to the NMR structure of ubiquitin (PDB ID 2MJB⁴⁷).

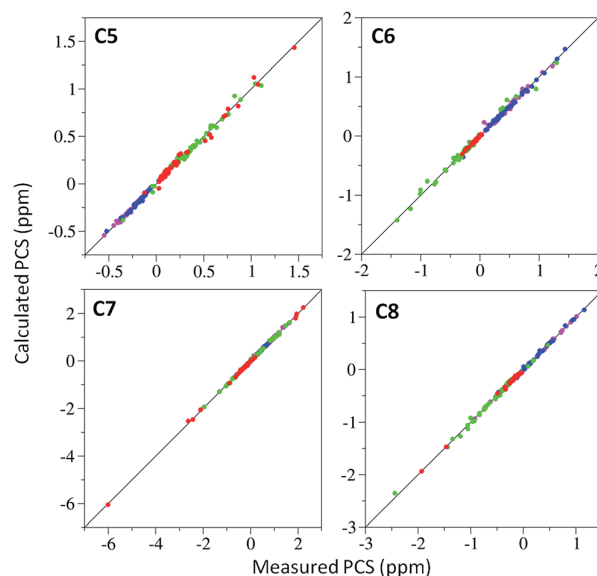


Fig. 3 Correlations between experimental and back-calculated PCSs for C5–C8 bound to ubiquitin A28C loaded with either Dy^{3+} (magenta), Tb^{3+} (blue), Tm^{3+} (green) or Yb^{3+} (red). Solid lines represent perfect correlation.

To investigate a potential change in the properties of different (S)-THP- Ln^{3+} tags with pH, we re-recorded the spectra of the C7- Ln^{3+} -tagged ubiquitin samples at pH 6.5 (Fig. S10 and S11†). Most notably, the PCSs of the Dy^{3+} -tagged sample were much smaller at pH 6.5 compared to pH 8 (slope 0.28, R^2 0.55). The majority of the Tb^{3+} PCSs were also reduced at lower pH, though to a lesser degree than for Dy^{3+} (slope 0.47, R^2 0.62). In contrast, the size of the PCSs observed in the Tm^{3+} (slope 0.95,

R^2 0.98) and Yb^{3+} (slope 0.94, R^2 0.99) samples was not significantly affected by the change in pH. The pH-dependence of the PCSs induced by each C7-Ln^{3+} complex likely reflects protonation/deprotonation processes involving the pendant arms and/or aquo ligands, leading to changes in the average charge and coordination geometry of each complex, and thus potentially their interactions with the protein surface and resulting metal ion positions. At both pHs, each C7-Ln^{3+} complex produced a single PCS for each affected nucleus, indicating that any processes such as protonation/deprotonation (and their effect on coordination geometry) are *fast* on the NMR timescale, thus the spectra are straightforward to interpret. The Tm^{3+} and Yb^{3+} complexes are likely to prove of most practical use over a wider, biologically relevant pH range.

For the data recorded at pH 8, we determined $\Delta\chi$ -tensors for lanthanide ions both individually, allowing independent metal ion positions (Table 1), and simultaneously with a common metal ion position for complexes of a given tag (Table S4†). In some cases there were significant differences between the individually and simultaneously determined $\Delta\chi$ -tensors, with the $\Delta\chi_{\text{ax}}$ component varying by up to 49% for the most extreme example of the C5-Dy^{3+} complex. Despite this, the Q -values of $\Delta\chi$ -tensors determined from either method were very respectable (0.02–0.12), demonstrating that the tensors from either approach are suitable for structural investigations. Individually derived $\Delta\chi$ -tensors produced Q -values that were universally lower than those of the simultaneously calculated tensors; however, the individually determined metal ion positions were up to 6.2 Å apart for different complexes of the same tag. This observation of different metal ion positions in individual $\Delta\chi$ -tensor fits has been noted previously^{24,49} and in those cases was attributed to the uncertainty in determining the metal ion position during the fitting procedure, which can also depend on the coverage and distribution of the PCSs over the tensor “space”. Thus, a common metal ion position that satisfactorily describes the PCSs of each metal ion is often used to increase stability of the metal ion coordinate and tensor components during the fitting. In this case, the apparent different sensitivities of each metal complex to pH (influencing their average charge, coordination geometry and possible interactions with the protein surface) could be seen as justification for the use of individually determined metal ion positions and $\Delta\chi$ -tensors. It is worth noting again that mobility of the metal ion, for instance due to flexibility of the tag linker, results in averaging of PCSs. By fitting a single tensor to these averaged values, we are describing an “effective $\Delta\chi$ -tensor”. The metal ion coordinate associated with this tensor should not be interpreted as a definitive point at which the metal ion is statically located.⁴¹ Unless specified otherwise, figures and values presented herein were derived using $\Delta\chi$ -tensors corresponding to individual metal ion positions.

Comparison of performance of the new tags with C1

With few exceptions, the lanthanide complexes of the new tags produced $\Delta\chi$ -tensors with $\Delta\chi_{\text{ax}}$ components of similar or greater magnitude to those of the corresponding C1 tag

conjugated to the same ubiquitin mutant.²⁸ This is a particularly interesting and non-intuitive observation in the case of C5, as it suggests that any increase in mobility of the tag, due to the loss of the bulky phenyl amide pendants of C1, is compensated for by the altered coordination environment and ligand field associated with the alcohol pendants and/or changes in secondary interactions with the protein, allowing C5 to generate sizeable paramagnetic effects.

In contrast to the case for C1, for which each lanthanide complex reliably produces PCSs of a predictable relative size and sign for a given nuclear spin (e.g. Tm^{3+} and Tb^{3+} PCSs are generally opposite in sign, with Tb^{3+} PCSs slightly larger in size), the relative order and size of PCSs induced by the new tags loaded with different lanthanide ions was quite variable (Fig. S12†). Correspondingly, the determined Euler angles of the $\Delta\chi$ -tensors from metal complexes of the same tag also varied to a larger extent than those observed for C1 (Fig. S13†), suggesting changes in coordination environment with each lanthanide ion, as alluded to above. The noted change in the orientation of the $\Delta\chi$ -tensor for each metal complex of the same tag is potentially a useful property, which can help resolve the redundant solutions that can be encountered in studies using PCSs (associated with the symmetry of the $\Delta\chi$ -tensor), without requiring multiple tagging sites or tags.⁵⁰

Initially, only the (*S*)-enantiomer of each tag was synthesised and assessed. However, given the large $\Delta\chi$ -tensors and excellent fits observed for C7, its enantiomer C8 and the corresponding C8-Ln^{3+} complexes were also synthesised and conjugated to ubiquitin A28C. Despite the same coordination environment of the lanthanide ions in complexes of either tag enantiomer, different PCSs, $\Delta\chi$ -tensors and metal ion positions (Fig. 2, 3, S8 and S12; Tables 1, S3 and S4†) were observed, likely due to the differences in their interaction with the chiral protein surface, arising from the opposite stereochemistry of the pendant arms. On average, the $\Delta\chi_{\text{ax}}$ components of each complex of the C7 and C8 tags were larger than those of the C1, C5 or C6 tags on ubiquitin A28C, suggesting that the very short linker is key to the tags' superior paramagnetic effects. A temperature titration of the C7-Tm^{3+} tagged sample showed no signs of additional cross-peaks due to conformational exchange (Fig. S14†), although at higher temperatures the observed PCSs were smaller, presumably due to increased tag mobility.

In order to investigate the rigidity of the C7 and C8 tags and their ability to induce partial alignment in the magnetic field, one bond ^1H – ^{15}N RDCs ($^1D_{\text{HN}}$) of the Tm^{3+} complexes of C7 and C8 were measured relative to the Y^{3+} -tagged protein. $^1D_{\text{HN}}$ RDCs up to 12.5 and 6.1 Hz were observed at 600 MHz for C7 and C8 respectively. Alignment tensors were determined by fitting the measured RDCs (Table S5†) to a structure of ubiquitin using single value decomposition within PALES⁵¹ (Table S6†). The measured and calculated $^1D_{\text{HN}}$ RDCs (Fig. 4A and B) are in good agreement and the principal axes of the alignment (Fig. 4C and D) and $\Delta\chi$ (Fig. 4E and F) tensors are very similar, demonstrating that the orientation of the tensors are relatively well defined for either enantiomer. The $\Delta\chi$ -tensor components derived from the alignment tensor parameters match very favourably with the PCS derived $\Delta\chi$ -tensor values for C7-Tm^{3+}



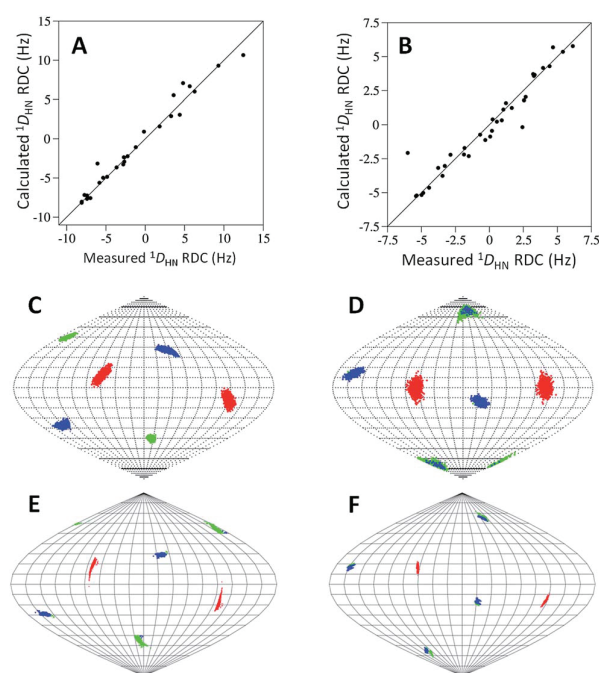


Fig. 4 Correlations between experimental and calculated $^1D_{\text{HN}}$ RDCs recorded at a ^1H NMR frequency of 600 MHz for **C7-Tm $^{3+}$** (A) and **C8-Tm $^{3+}$** (B) tagged ubiquitin A28C. Solid lines represent perfect correlation. Orientations of the principal axes of the alignment (C **C7**, D **C8**) and $\Delta\chi$ (E **C7**, F **C8**) tensors. The points show where the principal axes of the tensors penetrate the sphere with the axes coloured as follows: z (blue), y (green), x (red). For the alignment tensors, 1000 replicates of SVD calculation using the structural noise Monte-Carlo method ('-mcStruc') within PALES are shown. For the $\Delta\chi$ -tensors, 1000 replicates with a random 10% of the PCS data removed each time are shown. The convention $|z| > |y| > |x|$ is used to name the axes, resulting in swapping of the $|z|$ and $|y|$ axes in different fits when their magnitudes are similar.

(Tables 1 and S6; eqn S1†). However, for **C8-Tm $^{3+}$** , the alignment tensor predicted $\Delta\chi_{\text{ax}}$ and $\Delta\chi_{\text{rh}}$ are 62% and 54% of their respective PCS determined values, suggesting some degree of mobility is still present. It is not uncommon for alignment tensors to be smaller than $\Delta\chi$ -tensors.^{27,49,52} This is partly attributed to the greater sensitivity of RDCs to protein and tag movements than PCSs. The Q -factors of the alignment tensors are larger than those of the $\Delta\chi$ -tensors. Due to the chiral nature of the tags, **C7** may be engaged in different secondary

interactions with the protein, helping to limit its mobility to a greater degree than **C8**.

Further testing of **C7** and **C8** on a cysteine-bearing mutant of **GB1** and **HPPK**

To demonstrate the general utility of the **C7** and **C8** tags, lanthanide complexes of both tags were conjugated to a **GB1** Q32C mutant. Each sample produced a single set of PCSs from which $\Delta\chi$ -tensors were determined (Fig. S15 and S16; Tables 2, S7 and S8†). Differences between individual and simultaneously derived $\Delta\chi$ -tensors were apparent, though to a lesser extent than observed for ubiquitin. Both **Tb $^{3+}$** complexes, but particularly the **C7** complex, resulted in only small PCSs and significantly smaller $\Delta\chi$ -tensors on **GB1** compared to ubiquitin. Thus, the **GB1** spectra, which were recorded at pH 6.5, seem consistent with the ubiquitin spectra recorded at pH 6.5, in that they suggest that the **Tb $^{3+}$** (and likely **Dy $^{3+}$**) complexes are of less practical use at a lower pH. Conversely, both **Tm $^{3+}$** complexes resulted in sizable PCSs and $\Delta\chi$ -tensors. The $\Delta\chi_{\text{ax}}$ components of **C7-Tm $^{3+}$** and **C8-Tm $^{3+}$** on **GB1** Q32C are 73% and 91% of their respective values on ubiquitin A28C, demonstrating the influence of the tagging site and protein environment on the tags' performance.

This variability was further observed in our investigations of the 20 kDa-sized protein **HPPK** (to be fully reported elsewhere). Tagging at different sites produced $\Delta\chi$ -tensors with varied $\Delta\chi_{\text{ax}}$ components, up to $54.5 \times 10^{-32} \text{ m}^3$ for a **HPPK** S112C/C80A mutant tagged with **C7-Tm $^{3+}$** (Fig. S16 and S17; Tables 2 and S9†). Given that the PCS and RDC data for **C7-Tm $^{3+}$** tagged ubiquitin A28C had previously indicated that the chelate was relatively rigid on ubiquitin, such an increase in the $\Delta\chi_{\text{ax}}$ component for **HPPK** S112C/C80A was highly unexpected. Spectra of both proteins were recorded at pH 8, thus different deprotonation/protonation equilibria based on solvent water alone are insufficient to explain such variance. However, different interactions with the protein surface could also affect the charged state of the tag. For this particular **HPPK** mutant, the cysteine for tagging was introduced on the β -sheet of a short β -hairpin, which features an aspartic acid (**D107**) on the adjacent β -sheet. The calculated metal ion position is above and between **D107** and **S112**, which both point in the same direction in the **HPPK** crystal structure (Fig. S18†). The carboxyl group of **D107** could conceivably be interacting with either the hydroxyl pendants of the tag or directly with the lanthanide ion to

Table 2 $\Delta\chi$ -Tensor parameters for **C7** and **C8** tagged **GB1** Q32C and **C7** tagged **HPPK** S112C/C80A^{a,b}

Protein	Tag	Ln $^{3+}$	# PCS	$\Delta\chi_{\text{ax}}$	$\Delta\chi_{\text{rh}}$	Q	x	y	z	α	β	γ
GB1	C7	Tb$^{3+}$	47	2.2 (1.1)	1.0 (0.7)	0.09	29.244	29.993	13.297	18	39	60
		Tm$^{3+}$	37	−14.2 (0.9)	−4.4 (0.5)	0.06	31.745	29.577	12.618	145	56	80
	C8	Tb$^{3+}$	40	6.1 (0.4)	3.1 (0.3)	0.05	33.668	30.260	14.419	155	41	175
		Tm$^{3+}$	40	−14.9 (0.4)	−6.4 (0.6)	0.06	34.227	32.261	17.019	172	73	173
HPPK	C7	Tm$^{3+}$	81	54.5 (0.5)	12.5 (0.5)	0.04	14.304	13.802	13.906	149	55	127

^a See footnote a in Table 1. ^b Metal ion coordinates (x, y, z) for each tag are relative to the crystal structures of **GB1** (PDB ID 1PGA)⁵⁴ or **HPPK** (PDB ID 3QBC).⁴⁵



influence the charge of the chelate and its paramagnetic properties. In addition, HPPK samples tagged with C7 and C8 appeared more stable to precipitation than those tagged with C1 or C2, allowing the acquisition of multiple NMR experiments of each sample.

Experimental

Materials and methods

(1*S*,4*S*,7*S*)-1,4,7-Tris(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane,⁴⁶ 2-chloro-*N*-(2-(pyridin-2-ylidisulfanyl)ethyl)acetamide²⁸ and dimethyl 4-(hydroxymethyl)pyridine-2,6-dicarboxylate²³ were prepared following literature procedures. The synthesis of C5 and C6 is described in the ESI.[†]

Synthetic procedures

(*S*)-1-Chloro-3-(tritylthio)propan-2-ol (10). Triphenylmethanethiol (2.242 g, 8.11 mmol) was added to a solution of (*S*)-epichlorohydrin (500 mg, 5.40 mmol) and potassium fluoride (628 mg, 10.81 mmol) in MeOH (15 mL) and the mixture was stirred vigorously at room temperature for 72 h. Insoluble material was removed by filtration and the filtrate concentrated under reduced pressure. The resulting residue was washed with H₂O (10 mL) and Et₂O (10 mL) and the aqueous layer washed twice more with Et₂O (10 mL each). The organic layers were combined, dried with anhydrous MgSO₄ and concentrated under reduced pressure. The resulting oil was purified by silica flash chromatography (10% EtOAc in PET Spirits) to yield **10** as a colourless oil. Yield: 1.888 g (95%). ¹H NMR (400 MHz, MeOD) δ 7.41 (m, 6H), 7.29 (m, 6H), 7.22 (m, 3H), 3.41 (m, 3H, CHOH, CH₂Cl), 2.41 (m, 2H, CH₂S). ¹³C NMR (101 MHz, MeOD) δ 146.13 (C), 130.78 (CH), 128.94 (CH), 127.84 (CH), 71.33 (CHOH), 67.62 (C(Ph)₃), 49.032 (CH₂Cl), 36.96 (CH₂S). *R*_f (10% EtOAc in PET Spirits): 0.19.

(2*S*,2'*S*,2''*S*)-1,1',1''-(10-((*R*)-2-Hydroxy-3-(tritylthio)propyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triyl)tris(propan-2-ol) (11). Potassium carbonate (601 mg, 4.35 mmol) was added to a solution of (1*S*,4*S*,7*S*)-1,4,7-tris(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane (300 mg, 0.87 mmol) and **10** (321 mg, 0.87 mmol) in ACN (5 mL). The mixture was heated to reflux for 20 h, after which an additional equivalent of **10** (321 mg, 0.87 mmol) was added and refluxed for a further 4 h. After cooling to room temperature, insoluble salts were removed by filtration and the filtrate concentrated under reduced pressure. 1 M NaOH (25 mL) was added to the residue and washed with CHCl₃ (3 \times 25 mL). The organic layers were combined, dried with anhydrous MgSO₄ and concentrated under reduced pressure. The resulting residue was purified by silica flash chromatography (0–10% MeOH, 1% NH₃ in CHCl₃) to yield **11** as a yellow oil. Yield: 357 mg (60%). ¹H NMR (400 MHz, CDCl₃) δ 7.41 (m, 6H), 7.26 (m, 6H), 7.18 (m, 3H), 5.28 (br, 1H), 4.97 (br, 1H), 3.86 (m, 3H, CHCH₃), 3.46 (m, 1H, CHCH₂S), 2.95–2.73 (m, 8H), 2.45 (dd, *J* = 12.4, 5.2 Hz, 1H), 2.37–2.16 (m, 6H), 2.11 (dd, *J* = 4.5, 1.6 Hz, 1H), 2.08 (dd, *J* = 4.4, 1.7 Hz, 1H), 2.01 (m, 7H), 1.96–1.89 (m, 2H), 1.08 (d, *J* = 6.3 Hz, 3H, CH₃), 1.06 (d, *J* = 6.2 Hz, 6H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 144.93 (C), 129.76 (CH), 127.93 (CH), 126.64 (CH), 66.72 (C(Ph)₃), 66.19 (CHOH), 63.72 (CH₂),

63.12 (CH₂), 63.02 (CHOH), 62.86 (CHOH), 61.87, 51.50, 51.44, 51.11, 51.00 (previous 5 signals CH₂), 36.30 (CH₂S), 20.09 (CH₃), 20.00 (CH₃). LC-MS: *m/z* (ESI, 20 V) 436.3 (100%) [M + 2H-trityl]⁺, 679.4 (28%) [M + H]⁺.

(2*S*,2'*S*,2''*S*)-1,1',1''-(10-((*R*)-2-Hydroxy-3-(pyridin-2-ylidisulfanyl)propyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triyl)tris(propan-2-ol), trifluoroacetate salt (C7). Trifluoroacetic acid (1 mL) was added slowly to a solution of **11** (350 mg, 0.52 mmol) and triethylsilane (124 μ L, 0.77 mmol) in DCM (2 mL), forming a cloudy mixture that was stirred at room temperature for 1 h. Volatile reagents were removed by blowing a stream of N₂ over the open reaction vessel, before further concentrating under reduced pressure. The resulting residue was dissolved in MeOH (5 mL) and DCM (1 mL), before 2,2'-dipyridyldisulfide (229 mg, 1.04 mmol) was added and the solution stirred at room temperature for 15 min before concentrating under reduced pressure. The residue was washed between 0.1% TFA in H₂O (15 mL) and DCM (15 mL) and the aqueous layer purified by reverse-phase HPLC (0.1% TFA and a 5–100% ACN gradient over 20 min on a C18 preparative column). Fractions containing pure product were lyophilised to yield the trifluoroacetate salt of C7 as a yellow oil. Yield: 208 mg (39%, assuming a pentatrifluoroacetate salt). ¹H NMR (400 MHz, D₂O) δ 8.56 (m, 1H, *H*₆ of Pyr), 8.30 (ddd, *J* = 8.4, 7.7, 1.6 Hz, 1H, *H*₄ of Pyr), 8.13 (d, *J* = 8.4 Hz, 1H, *H*₃ of Pyr), 7.70 (m, 1H, *H*₅ of Pyr), 4.25–4.04 (m, 4H, CHOH), 3.62–3.42 (m, 4H), 3.34–2.96 (m, 14H), 2.90–2.55 (m, 8H), 1.14 (d, *J* = 6.1 Hz, 6H, CH₃), 1.08 (d, *J* = 6.3 Hz, 3H, CH₃). ¹³C NMR (101 MHz, D₂O) δ 155.55 (C2 of Pyr), 145.40 (C4 of Pyr), 142.73 (C6 of Pyr), 125.44 (C3 of Pyr), 124.03 (C5 of Pyr), 64.94 (CHOH), 62.71 (CHOH), 60.82 (CHOH), 59.82, 59.74, 56.75, 50.49, 50.35, 49.76, 49.27 (previous 7 signals CH₂), 43.51 (CH₂S), 20.27 (CH₃), 19.78 (CH₃), 19.75 (CH₃). HRMS (ESI) *m/z* calcd [M + H]⁺ C₂₅H₄₈N₅O₄S₂: 546.3142, found: 546.3140. Analytical HPLC: *t*_R 4.21 min, 98% (254 nm).

Formation of lanthanide complexes

Complexes of C5 and C6 were prepared by refluxing the ligands for 18 h with 2 equivalents of Y³⁺, Dy³⁺, Tb³⁺, Tm³⁺ or Yb³⁺-trichloride salts in a 1 : 1 ACN : H₂O solution adjusted to neutral pH, followed by purification by HPLC (0.1% TFA and a 0–80% ACN gradient on a C18 preparative column) to afford the complexes as off-white solids after lyophilisation. In the case of C6, TCEP was added prior to purification to prevent disulfide formation.

Complexes of C7 and C8 were most readily prepared from **11**. An example of the formation of the C7-Yb³⁺ complex follows. A solution of **11** (30 mg, 0.044 mmol) and YbCl₃ (19 mg, 0.066 mmol) in MeOH (1.5 mL) was heated at 50 °C for 4 h, after which LCMS analysis indicated complete complexation. The solution was cooled to room temperature, then 2,2'-dipyridyldisulfide (29 mg, 0.13 mmol) and silver nitrate (37 mg, 0.22 mmol) added whilst vigorously stirring, forming a milky beige mixture, before formation of a beige precipitate that eventually turned grey. After 2 h, LCMS analysis indicated complete trityl deprotection and thiol activation, and the mixture was concentrated under reduced pressure. 0.1% TFA in H₂O (5 mL) and DCM (5 mL) were added to the grey residue and the suspension transferred



to a 15 mL centrifuge tube. The suspension was shaken vigorously and the precipitate sedimented and organic and aqueous phases separated, by centrifugation for 3 min at 2000 rcf. The aqueous phase was carefully removed and purified by reverse-phase HPLC (0.1% TFA and a 5–100% ACN gradient over 30 min on a C18 preparative column). Fractions containing pure product were lyophilised to yield the trifluoroacetate salt of **C7-Yb**³⁺ as an off-white solid. Yield: 17 mg (34%, assuming a tetratrifluoroacetate salt).

Working stock solutions of each metal complex were prepared at 20 mM in H₂O and stored frozen at –20 °C when not in use.

C5-Y³⁺. ¹H NMR (400 MHz, D₂O) δ 8.56 (m, 1H), 8.18 (m, 1H), 8.08 (m, 1H), 7.60 (ddd, J = 7.5, 5.5, 1.1 Hz, 1H), 4.71 (m, 1H), 4.56 (m, 2H), 4.03 (d, J = 16.3 Hz, 1H), 3.94 (m, 1H), 3.61–3.29 (m, 9H), 3.21–3.04 (m, 4H), 2.82 (m, 1H), 2.74–2.28 (m, 12H), 1.36 (d, J = 5.9 Hz, 3H), 1.28 (d, J = 5.9 Hz, 3H), 1.24 (d, J = 5.8 Hz, 3H). HRMS (ESI) m/z calcd [M – 2H]⁺ C₂₆H₄₆N₆O₄S₂Y: 659.2091, found: 659.2087.

C6-Y³⁺. ¹H NMR (400 MHz, D₂O) δ 8.02 (s, 1H), 7.76 (s, 1H), 4.55–4.34 (m, 3H), 3.91 (s, 2H), 3.81 (d, J = 15.0 Hz, 1H), 3.73–3.46 (m, 7H), 3.24–3.05 (m, 4H), 2.66 (d, J = 13.1 Hz, 2H), 2.58–2.35 (m, 6H), 2.32–2.15 (m, 4H), 1.43 (d, J = 5.8 Hz, 3H), 1.24 (d, J = 5.7 Hz, 3H), 0.88 (d, J = 5.7 Hz, 3H). HRMS (ESI) m/z calcd [M – 2H]⁺ C₂₂H₄₃N₅O₅SY: 614.2043, found: 614.2046.

C7-Y³⁺. ¹H NMR (400 MHz, D₂O) δ 8.38 (t, J = 4.5 Hz, 1H), 7.90–7.74 (m, 2H), 7.31 (m, 1H), 4.61–4.44 (m, 2H), 4.21–4.01 (m, 2H), 3.59–2.57 (m, 20H), 2.44–2.10 (m, 6H), 1.20 (m, 5H), 1.13 (d, J = 6.1 Hz, 2H), 1.08 (d, J = 6.1 Hz, 2H). HRMS (ESI) m/z calcd [M – 2H]⁺ C₂₅H₄₅N₅O₄S₂Y: 632.1917, found: 632.1963.

¹H NMR spectra and HRMS of the Yb³⁺ complexes of **C5–C7**, are shown in Fig. S1–S6.†

NMR sample preparation

Uniformly ¹⁵N-labelled human ubiquitin A28C was prepared as described.²⁴ Prior to tagging the protein was first reduced by stirring with a 10-fold excess of DTT for 1 h at room temperature, before passage over a PD-10 column equilibrated with degassed buffer (50 mM HEPES, pH 8).

For the **C5** and **C7/8** tags a 5-fold excess of the respective lanthanide complex was added to a solution of protein and stirred at room temperature for 2 h. Excess tag was removed by passage over a PD-10 column before the sample was concentrated using a Millipore ultrafilter (3 kDa) to a final protein concentration of approximately 100 μ M.

In order to tag **C6**, the protein cysteines were first pre-activated by addition of 10 equivalents of 5,5'-dithiobis-2-nitrobenzoic acid (DTNB), producing a yellow coloured solution that was allowed to stir at room temperature for 1 h. Excess DTNB and TNB^{2–} leaving group were removed by passage through a PD10 column, yielding a colourless solution. A 5-fold excess of the respective **C6** complex was then added, forming a yellow solution that was stirred at room temperature for 2 h. Excess tag and TNB^{2–} leaving group was removed by passage over a PD-10 column and samples were concentrated as above.

NMR spectroscopy

Spectra of differently tagged ubiquitin A28C in 90%/10% H₂O/D₂O, 50 mM HEPES, pH 8.0, were recorded at 25 °C on either Varian INOVA or Bruker Avance 600 MHz NMR spectrometers equipped with cryogenic probes. ¹H^N PCSs and ¹D_{HN} couplings were measured by recording ¹⁵N-fast-HSQC spectra with and without the 180° (¹H) pulse during the ¹⁵N (t_1) evolution time. 2D ¹⁵N-fast-HSQC were typically acquired with $t_{1\max}$ (¹⁵N) = 51–62 ms and $t_{2\max}$ (¹H) = 142 ms.

Calculation of $\Delta\chi$ and alignment tensors

Fitting of $\Delta\chi$ -tensors was carried out within the program Numbat.⁵³ The tensors were fitted to the first conformer of the NMR structure of ubiquitin (PDB 2MJB⁴⁷). Unambiguous PCS assignments were used to calculate an initial estimate of the $\Delta\chi$ -tensor, from which iterative cycles of further assignment and recalculation were made. The $\Delta\chi$ -tensors for GB1 and HPPK were determined in an analogous way, fitting to the crystal structures of GB1 (PDB 1PGA⁵⁴) and HPPK (PDB 3QBC⁴⁵).

Backbone amide ¹D_{HN} RDCs were fitted to the first conformer of the NMR structure of ubiquitin (PDB 2MJB⁴⁷) using single value decomposition *via* the “bestFit” flag in PALES.⁵¹

Conclusions

We have presented the synthesis of three new LBT designs. Each tag is capable of binding lanthanide ions tightly and producing significant PCSs without need for the addition of free paramagnetic metal ions to protein samples. Each design features hydroxypropyl pendant arms, rendering the tags smaller and more hydrophilic than previously reported DOTA-style tags. The **C5** tag can be readily synthesised and displayed comparable paramagnetic effects to **C1**, whose utility has been proven in several studies.^{55–57} The **C6** tag also performed comparably with **C1** on ubiquitin, however it features the longest synthesis of any of the tags and requires more protein handling *via* DTNB activation for conjugation. The **C7/8** design features a particularly short linker, resulting in limited mobility relative to the protein surface, hence generating the largest paramagnetic effects on ubiquitin. The capability of **C7** and **C8** to produce paramagnetic effects on other proteins was further demonstrated on GB1 and HPPK. The performance of the tags varied with factors including the lanthanide used, pH and site of conjugation. Given their favourable properties, it is anticipated that **C7** and **C8** (particularly their Tm³⁺ complexes) will prove useful in the investigation of a wide range of biologically interesting proteins by paramagnetic NMR spectroscopy.

Acknowledgements

An Australian Research Council Future Fellowship to BG (FT130100838) and Discovery Grants to GO and BG (DP120100561 and DP150100383) are gratefully acknowledged.



Notes and references

- 1 G. Otting, *Annu. Rev. Biophys.*, 2010, **39**, 387–405.
- 2 J. Koehler and J. Meiler, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **59**, 360–389.
- 3 W.-M. Liu, M. Overhand and M. Ubbink, *Coord. Chem. Rev.*, 2014, **273–274**, 2–12.
- 4 M. Allegrozzi, I. Bertini, M. B. L. Janik, Y. M. Lee, G. Liu and C. Luchinat, *J. Am. Chem. Soc.*, 2000, **122**, 4154–4161.
- 5 L. Lee and B. D. Sykes, *Biophys. J.*, 1980, **32**, 193–210.
- 6 I. Baig, I. Bertini, C. Del Bianco, Y. K. Gupta, Y.-M. Lee, C. Luchinat and A. Quattrone, *Biochemistry*, 2004, **43**, 5562–5573.
- 7 I. Bertini, Y. K. Gupta, C. Luchinat, G. Parigi, M. Peana, L. Sgheri and J. Yuan, *J. Am. Chem. Soc.*, 2007, **129**, 12786–12794.
- 8 S. Dasgupta, X. Hu, P. H. J. Keizers, W.-M. Liu, C. Luchinat, M. Nagulapalli, M. Overhand, G. Parigi, L. Sgheri and M. Ubbink, *J. Biomol. NMR*, 2011, **51**, 253–263.
- 9 C. Luchinat, M. Nagulapalli, G. Parigi and L. Sgheri, *J. Magn. Reson.*, 2012, **215**, 85–93.
- 10 G. Pintacuda, A. Y. Park, M. A. Keniry, N. E. Dixon and G. Otting, *J. Am. Chem. Soc.*, 2006, **128**, 3696–3702.
- 11 T. Saio, M. Yokochi, H. Kumeta and F. Inagaki, *J. Biomol. NMR*, 2010, **46**, 271–280.
- 12 P. H. J. Keizers, B. Mersinli, W. Reinle, J. Donauer, Y. Hiruma, F. Hannemann, M. Overhand, R. Bernhardt and M. Ubbink, *Biochemistry*, 2010, **49**, 6846–6855.
- 13 Y. Hiruma, M. A. S. Hass, Y. Kikui, W. M. Liu, B. Olmez, S. P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Löhr, H. Schwalbe, M. Nojiri and M. Ubbink, *J. Mol. Biol.*, 2013, **425**, 4353–4365.
- 14 M. John, G. Pintacuda, A. Y. Park, N. E. Dixon and G. Otting, *J. Am. Chem. Soc.*, 2006, **128**, 12910–12916.
- 15 T. Zhuang, H. Lee, B. Imperiali and J. H. Prestegard, *Protein Sci.*, 2008, **17**, 1220–1231.
- 16 T. Saio, K. Ogura, K. Shimizu, M. Yokochi, T. R. Burke and F. Inagaki, *J. Biomol. NMR*, 2011, **51**, 395–408.
- 17 J.-Y. Guan, P. H. J. Keizers, W.-M. Liu, F. Loehr, S. P. Skinner, E. A. Heeneman, H. Schwalbe, M. Ubbink and G. D. Siegal, *J. Am. Chem. Soc.*, 2013, **135**, 5859–5868.
- 18 C. Schmitz, R. Vernon, G. Otting, D. Baker and T. Huber, *J. Mol. Biol.*, 2012, **416**, 668–677.
- 19 H. Yagi, K. B. Pilla, A. Maleckis, B. Graham, T. Huber and G. Otting, *Structure*, 2013, **21**, 883–890.
- 20 X.-C. Su and G. Otting, *J. Biomol. NMR*, 2010, **46**, 101–112.
- 21 X.-C. Su, K. McAndrew, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2008, **130**, 1681–1687.
- 22 K. Barthelmes, A. M. Reynolds, E. Peisach, H. R. A. Jonker, N. J. DeNunzio, K. N. Allen, B. Imperiali and H. Schwalbe, *J. Am. Chem. Soc.*, 2011, **133**, 808–809.
- 23 X.-C. Su, B. Man, S. Beeren, H. Liang, S. Simonsen, C. Schmitz, T. Huber, B. A. Messerle and G. Otting, *J. Am. Chem. Soc.*, 2008, **130**, 10486–10487.
- 24 J. D. Swarbrick, P. Ung, S. Chhabra and B. Graham, *Angew. Chem., Int. Ed.*, 2011, **50**, 4403–4406.
- 25 H. Yagi, A. Maleckis and G. Otting, *J. Biomol. NMR*, 2013, **55**, 157–166.
- 26 P. H. J. Keizers, J. F. Desreux, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2007, **129**, 9292–9293.
- 27 T. Saio, K. Ogura, M. Yokochi, Y. Kobashigawa and F. Inagaki, *J. Biomol. NMR*, 2009, **44**, 157–166.
- 28 B. Graham, C. T. Loh, J. D. Swarbrick, P. Ung, J. Shin, H. Yagi, X. Jia, S. Chhabra, N. Barlow, G. Pintacuda, T. Huber and G. Otting, *Bioconjugate Chem.*, 2011, **22**, 2118–2125.
- 29 T. Ikegami, L. Verdier, P. Sakhaii, S. Grimme, B. Pescatore, K. Saxena, K. M. Fiebig and C. Griesinger, *J. Biomol. NMR*, 2004, **29**, 339–349.
- 30 G. Pintacuda, A. Moshref, A. Leonchiks, A. Sharipo and G. Otting, *J. Biomol. NMR*, 2004, **29**, 351–361.
- 31 M. D. Vlasie, C. Comuzzi, A. M. C. H. van den Nieuwendijk, M. Prudêncio, M. Overhand and M. Ubbink, *Chem. – Eur. J.*, 2007, **13**, 1715–1723.
- 32 P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**, 14802–14812.
- 33 D. Häussinger, J. Huang and S. Grzesiek, *J. Am. Chem. Soc.*, 2009, **131**, 14761–14767.
- 34 W. P. Cacheris, S. K. Nickle and A. D. Sherry, *Inorg. Chem.*, 1987, **26**, 958–960.
- 35 S. Aime, M. Botta, M. Fasano, M. P. M. Marques, C. F. G. C. Geraldles, D. Pubanz and E. Merbach, *Inorg. Chem.*, 1997, **1669**, 2059–2068.
- 36 W.-M. Liu, S. P. Skinner, M. Timmer, A. Blok, M. A. S. Hass, D. V. Filippov, M. Overhand and M. Ubbink, *Chem. – Eur. J.*, 2014, **20**, 6256–6528.
- 37 C. T. Loh, K. Ozawa, K. L. Tuck, N. Barlow, T. Huber, G. Otting and B. Graham, *Bioconjugate Chem.*, 2013, **24**, 260–268.
- 38 K. A. Chin, J. R. Morrow, C. H. Lake and M. R. Churchill, *Inorg. Chem.*, 1994, **33**, 656–664.
- 39 M. Lelli, G. Pintacuda, A. Cuzzola and L. Di Bari, *Chirality*, 2005, **17**, 201–211.
- 40 C.-H. Huang and J. R. Morrow, *Inorg. Chem.*, 2009, **48**, 7237–7243.
- 41 D. Shishmarev and G. Otting, *J. Biomol. NMR*, 2013, **56**, 203–216.
- 42 B. Man, X.-C. Su, H. Liang, S. Simonsen, T. Huber, B. A. Messerle and G. Otting, *Chem. – Eur. J.*, 2010, **16**, 3827–3832.
- 43 X. Jia, A. Maleckis, T. Huber and G. Otting, *Chem. – Eur. J.*, 2011, **17**, 6830–6836.
- 44 Q.-F. Li, Y. Yang, A. Maleckis, G. Otting and X.-C. Su, *Chem. Commun.*, 2012, **48**, 2704–2706.
- 45 S. Chhabra, O. Dolezal, B. M. Collins, J. Newman, J. S. Simpson, I. G. Macreadie, R. Fernley, T. S. Peat and J. D. Swarbrick, *PLoS One*, 2012, **7**, e29444.
- 46 J. Hammell, L. Buttarazzi, C.-H. Huang and J. R. Morrow, *Inorg. Chem.*, 2011, **50**, 4857–4867.
- 47 A. S. Maltsev, A. Grishaev, J. Roche, M. Zasloff and A. Bax, *J. Am. Chem. Soc.*, 2014, **136**, 3752–3755.
- 48 K. Chin and J. Morrow, *Inorg. Chem.*, 1994, **33**, 5036–5041.

- 49 F. Peters, M. Maestre-Martinez, A. Leonov, L. Kovačič, S. Becker, R. Boelens and C. Griesinger, *J. Biomol. NMR*, 2011, **51**, 329–337.
- 50 Y. Kobashigawa, T. Saio, M. Ushio, M. Sekiguchi, M. Yokochi, K. Ogura and F. Inagaki, *J. Biomol. NMR*, 2012, **53**, 53–63.
- 51 M. Zweckstetter, *Nat. Protoc.*, 2008, **3**, 679–690.
- 52 Y. Yang, J. Wang, Y. Pei and X. Su, *Chem. Commun.*, 2015, **51**, 2824–2827.
- 53 C. Schmitz, M. J. Stanton-Cook, X.-C. Su, G. Otting and T. Huber, *J. Biomol. NMR*, 2008, **41**, 179–189.
- 54 T. Gallagher, P. Alexander, P. Bryan and G. Gilliland, *Biochemistry*, 1994, **33**, 4721–4729.
- 55 L. de la Cruz, T. H. D. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2011, **133**, 19205–19215.
- 56 H. Yagi, D. Banerjee, B. Graham, T. Huber, D. Goldfarb and G. Otting, *J. Am. Chem. Soc.*, 2011, **133**, 10418–10421.
- 57 K. Jaudzems, X. Jia, H. Yagi, D. Zhulenkova, B. Graham, G. Otting and E. Liepinsh, *J. Mol. Biol.*, 2012, **424**, 42–53.



Chemical Biology

Installation of a Rigid EDTA-Like Motif into a Protein α -Helix for Paramagnetic NMR Spectroscopy with Cobalt(II) IonsJames D. Swarbrick,* Phuc Ung, Matthew L. Dennis, Michael D. Lee, Sandeep Chhabra, and Bim Graham*^[a]

Abstract: Coupling two copies of an iminodiacetic acid–cysteine hybrid ligand to a pair of cysteine residues positioned in an $i, i+4$ arrangement within a protein α -helix leads to generation of an EDTA-like metal ion-binding motif. Rigid binding of a Co^{II} ion by this motif produces pseudo-contact shifts suitable for paramagnetic NMR structural studies.

The site-specific immobilisation of paramagnetic metal ions within the molecular framework of a protein induces a number of useful effects that can be measured by NMR spectroscopy, including pseudo-contact shifts (PCSs), paramagnetic relaxation enhancements (PREs), cross correlated relaxation and residual dipolar couplings (RDCs).^[1] Analysis of these NMR effects provides valuable restraints for the refinement of protein structures, determination of protein–protein and protein–ligand complexes,^[2] investigation of protein dynamics^[3] and de novo protein structure determinations.^[4] PCSs are particularly noteworthy as they are readily measured in sensitive 2D NMR experiments, offering potential for investigations of large or challenging macromolecular systems.^[5] They are observed as simple chemical shift changes (δ_{PCS}) between diamagnetic and paramagnetic samples and can provide long-range distance and angular information for each observable nuclear spin up to and beyond 40 Å from the metal centre. The magnitude and sign of a PCS for a nuclear spin can be calculated from knowledge of its position, given by the distance from the metal centre and the polar angles, θ and φ , with respect to the principal axes of the anisotropic magnetic susceptibility ($\Delta\chi$) tensor:

$$\delta_{\text{PCS}} = \frac{1}{12\pi r^3} [\Delta\chi_{\text{ax}}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{\text{rh}}\sin^2\theta\cos^2\varphi]$$

From the measurement of at least eight PCSs, and a structural model, the magnitude of the axial ($\Delta\chi_{\text{ax}}$) and rhombic ($\Delta\chi_{\text{rh}}$) components of the $\Delta\chi$ tensor can be determined, as well as its

rotation relative to the protein frame (given by the three Euler angles, α , β and γ). This gives access to a metal-centred coordinate system, allowing each nuclear spin to be interrogated in three-dimensional space by virtue of its measured PCS.

Typically, paramagnetic lanthanide(III) ions are employed to produce PCSs by a number of distinct methodologies including substitution of metal ions (e.g., Ca^{II} or Mg^{II}) in metalloproteins,^[2a,b] recombinant introduction of lanthanide-binding peptides (LBPs) into proteins,^[6] and the site-specific conjugation of synthetic lanthanide-binding tags (LBTs) onto the surface of proteins.^[1c,7] Although less commonly utilised, high-spin Co^{II} complexes can also generate useful hyperfine and PCS data. Substitution of Cu^{II} or Mg^{II} with Co^{II} has served to probe the active site structure of various cupredoxins^[8] and the structure of a DNA–drug complex.^[9] LBTs based on dipicolinic acid (DPA)^[10] as well as a genetically encoded bipyridyl-alanine (Bpy-Ala) amino acid^[11] have also been employed in combination with Co^{II} ions for paramagnetic NMR spectroscopy.

Recently, we reported two new LBTs derived from the hybridisation of L-cysteine with either iminodiacetic acid (IDA)^[12] or nitrilotriacetic acid (NTA).^[13] Attachment of these to proteins via disulfide bond formation produces lanthanide-binding motifs suitable for the generation of PCSs.^[4a,14] We now report that the installation of two copies of the IDA-based tag (1, Figure 1) at the $i, i+4$ positions of an α -helix can be used to generate a hexadentate EDTA-like metal ion-binding motif. Binding of a Co^{II} ion by this motif reveals a rigidly bound metal centre imparted by the two-point immobilisation strategy. Furthermore, the combination of high rigidity, the small PRE effect of the Co^{II} ion and α -helix attachment, mean that all tensor parameters, including the metal ion coordinate, can be accurately derived solely from the PCSs of the amides of the helix residues.

IDA is known to form 2:1 octahedral complexes with transition metal ions.^[15] Modelling studies (Figure S1 in the Support-



Figure 1. Structure of the IDA-based tagging agent, 1, and a schematic representation of Co^{II} ion chelation by two IDA ligands attached to a protein α -helix via cysteine residues.

[a] Dr. J. D. Swarbrick, Dr. P. Ung, M. L. Dennis, M. D. Lee, Dr. S. Chhabra, Dr. B. Graham
Monash Institute of Pharmaceutical Sciences, Monash University
381 Royal Parade, Parkville 3052, Victoria (Australia)

Supporting information and ORCID(s) from the author(s) for this article are available on the WWW under <http://dx.doi.org/10.1002/chem.201503139>.

ing Information) indicated that attachment of two copies of **1** to the $i, i+4$ positions of an α -helix could permit formation of such species without generating excess strain and sample limited conformational space translating to a relatively high degree of rigidity. To test this, two copies of the pyridine disulfide-activated tag, **1**, were conjugated to a pair of cysteines located on the α -helix of a mutant of human ubiquitin, UbqE24C/A28C. Zn^{II} ions were subsequently added in a sub-stoichiometric amount to obtain a diamagnetic reference spectrum. Assignment of the diamagnetic cross-peaks in the 2D ¹⁵N fast-HSQC spectrum (Figure 2) was facilitated by analysis of a HNCA spectrum. Only small chemical shift changes were observed after ligation of **1** and following the addition of Zn^{II} ions, and were mostly confined to the helix amides (Figure S2 in the Supporting Information). Removal of the Zn^{II} ions was achieved by addition of 1 mM EDTA and passage over a desalting column.

Addition of Co^{II} ions to the bis-IDA-tagged ubiquitin resulted in a single set of shifted resonances (Figure 2), indicative of a Co^{II}-protein complex with apparent chiral purity. Assignment of the paramagnetic NMR spectrum and derivation of the $\Delta\chi$ tensor was carried out within Numbat^[16] by fitting the PCSs to the NMR structure of human ubiquitin (2MJB).^[17] A low Q factor (5.3%) and a good correlation ($R^2=0.99$) between 66 measured ¹HN (for residues 1–71) and ten additional ¹⁵N PCSs (derived from residues along the helix) with their calculated values was found (Figure 3, Table S1 and Figure S3 in the Supporting Information), indicating correct assignments and that any change in the structure of ubiquitin caused by appendage of the Co^{II}-bis-IDA tag is minor. The metal ion position was found to lie on a line drawn perpendicular to the helix axis, bisecting the two C α atoms of the two cysteines and about approximately 7 Å from the helix (Figure 4A).

The main Z axis of the $\Delta\chi$ tensor is tilted slightly ($\approx 18^\circ$) from the plane normal to the helix axis, with the X axis pointing along the C α -C α ($i, i+4$) vector. An isosurface representation of the $\Delta\chi$ tensor (Figure 4B) maps regions in space with identical PCSs and reveals a tensor which is predominantly ax-

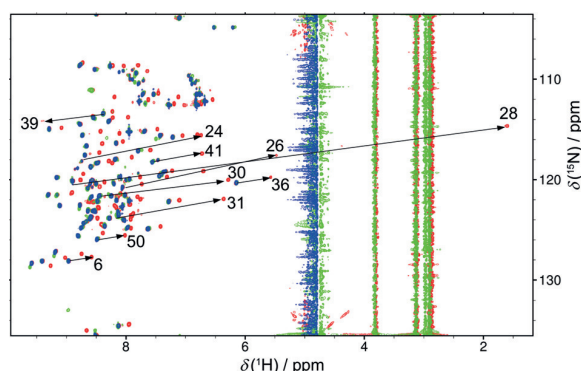


Figure 2. Superposition of ¹⁵N fast-HSQC spectra recorded on approximately 100 μ M solutions of ¹⁵N-labelled UbqE24C/A28C dissolved in 50 mM HEPES (pH 7.5), doubly tagged with **1** (blue) and with the addition of 60 μ M CoCl₂ (red) or 60 μ M ZnCl₂ (green). Arrows indicate PCSs of selected backbone amides.

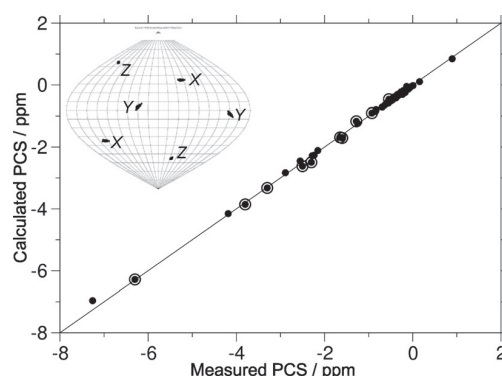


Figure 3. Correlation between observed and back-calculated ¹H PCSs from amides 1–71 (filled circles) and ¹⁵N PCSs from the helix (amides 22–34; open circles) for Co^{II}-loaded ¹⁵N-labelled UbqE24C/A28C doubly-tagged with **1**. Inset: Sanson–Flamsteed world map projections of the Monte Carlo error analysis, in which 10% of the data was randomly removed, showing the distribution of the principal axes of the $\Delta\chi$ tensor for 250 replicates.

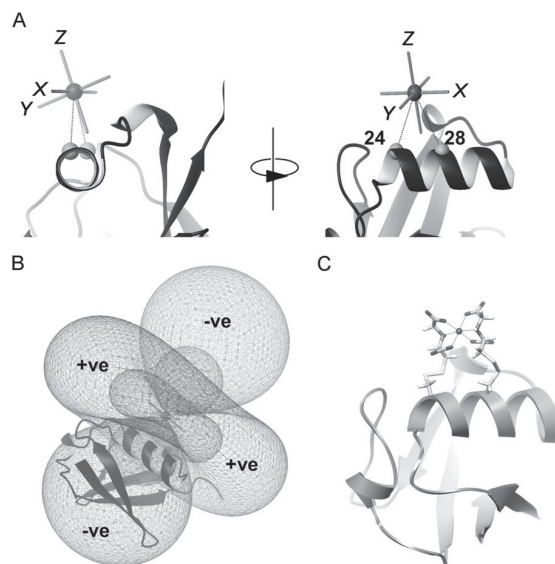


Figure 4. A) Position of Co^{II} (darker sphere) and orientation of the principal axes of the $\Delta\chi$ tensor relative to the C α atoms of Cys24 and Cys28 (light grey spheres). B) Isosurface representation of the $\Delta\chi$ tensor contoured at ± 1 ppm (inner lobes) and ± 0.1 ppm (outer lobes). C) Model of Co^{II}-bis-IDA-tagged UbqE24C/A28C complex.

ially symmetric, with the majority of the amides falling within the same lobe. Consequently, almost all the PCSs observed have a negative sign and the resonances shift upfield relative to the diamagnetic reference (Figure 2). The low scattering of the principal axes orientations as shown in the Sanson–Flamsteed projection in Figure 3, indicates that the orientation of the $\Delta\chi$ tensor is well defined by the PCS data. Moreover, removal of the large (and more difficult to assign) PCSs did not significantly affect the metal ion coordinate, the magnitude of the $\Delta\chi$ tensor or the associated errors (Table 1), indicating that the metal ion position and the $\Delta\chi$ tensor parameters are well

Table 1. $\Delta\chi$ tensor parameters.

	$\Delta\chi_{Ax}^{[c]}$	$\Delta\chi_{Rh}^{[c]}$	x	y	z
PCS ^[a] (error) ^[d]	$-7.3 (\pm 0.04)$	$-1.4 (\pm 0.06)$	$2.8 (\pm 0.05)$	$4.2 (\pm 0.07)$	$-12.6 (\pm 0.04)$
PCS ^[b] (error) ^[d]	$-7.3 (\pm 0.06)$	$-1.5 (\pm 0.08)$	$3.1 (\pm 0.09)$	$4.7 (\pm 0.11)$	$-12.2 (\pm 0.08)$

[a] 66 ^1H N and 10 ^{15}N PCSs included. [b] PCSs $> |2|$ ppm removed. [c] $\Delta\chi_{Ax}$ and $\Delta\chi_{Rh}$ are measured in 10^{-32} m^3 . The three Euler angles (α , β , γ) in UTR representation are 45° , 145° and 173° , respectively. [d] From 250 Monte Carlo replicants in which 10% of the PCS data was randomly removed.

defined by the PCS data and not dependent on the inclusion of the large PCSs.

Significantly, the metal ion coordinate (2.7 ± 0.2 , 4.2 ± 0.3 , -12.5 ± 0.2) and tensor parameters ($\Delta\chi_{Ax} = -6.8 \times 10^{-32} \text{ m}^3$, $\Delta\chi_{Rh} = -1.6 \times 10^{-32} \text{ m}^3$, $\alpha = 50^\circ$, $\beta = 144^\circ$, $\gamma = 8^\circ$) determined from just the PCSs of the helix residues (23–34) compared very favourably with the values derived from the full PCS dataset (Table 1, Figure S4 in the Supporting Information).

Tag rigidity is integral to tag performance and application. A rigid tag is paramount if dynamic analyses are to be successfully undertaken.^[3] To date, most tags are attached by a single tether and depending on the local protein environment are invariably mobile to some extent. While the PCSs obtained from partially mobile tags can still be accurately predicted, they are limited to nuclei located not too close to the metal ion and to those that reside within the protein PCS “sphere”^[18] which leads to poor prediction of PCSs in multiprotein complex assemblies. For Co^{II} -bis-IDA bound to ubiquitin, inclusion of the large ($> |2|$ ppm) ^1H N and ^{15}N PCSs from amides of the helix (Figure 3 and Figure S3) yielded excellent agreement with the calculated values, providing strong evidence that the Co^{II} ion is held rigidly within the protein frame and that there is little, if any, distortion of the helix. Furthermore, the PCS-calculated metal ion position was compatible with a physical model (Figure 4C and Figure S1) appended to ubiquitin, providing supporting evidence that the Co^{II} -bis-IDA tag has limited mobility.

The magnitude of the axial component of the $\Delta\chi$ tensor (Table 1) is also concordant with that predicted for a rigidly bound Co^{II} ion with octahedral geometry,^[19] and similar to that determined for Co^{II} -loaded azurin [$(6.8 \pm 0.3) \times 10^{-32} \text{ m}^3$].^[8a] It is also similar to that observed for the Co^{II} -loaded H87Bpy-Ala mutant of the West Nile Virus NS2B-NS3 protease [$(6.9 \pm 0.4) \times 10^{-32} \text{ m}^3$], which, based on the metal ion position, made an additional coordination to the carboxyl group of Asp145.^[11]

To confirm that the rigid binding of the Co^{II} ion is a consequence of coordination by both IDA ligands, we investigated the case of a single IDA-bound Co^{II} ion using two single-Cys ubiquitin mutants, UbiquA28C and UbiquE24A/A28C/D32A. UbiquA28C retains the native acidic residues E24 and D32 that could potentially assist with Co^{II} immobilisation, while UbiquE24A/A28C/D32A is devoid of acidic residues on the helix or within 8 Å of the Cys28 C β atom. For Co^{II} -IDA-tagged UbiquE24A/A28C/D32A, far smaller and fewer PCSs were observed than for Co^{II} -bis-IDA-tagged UbiquE24C/A28C (Figures S5, S6, Table S2 in the Supporting Information), as anticipated from the relatively exposed nature of the tagging site. The magnitude of the resulting $\Delta\chi$ tensor [$\Delta\chi_{Ax} = (3.3 \pm 2.0) \times 10^{-32} \text{ m}^3$, $\Delta\chi_{Rh} = (2.0 \pm 1.2) \times 10^{-32} \text{ m}^3$] is less than half that af-

forded by the bis-IDA chelate motif, and the large uncertainty in the tensor magnitude (ascribable to the low number of sizeable PCSs) indicates that a single IDA tagging approach is likely to be of little practical utility, unless the tagging site is sterically crowded or there are additional coordinating residues to assist with “anchoring” the Co^{II} ion in place. Even with such residues present, the Co^{II} -IDA-tagged UbiquA28C mutant yielded spectra that were notably inferior to that of Co^{II} -bis-IDA-tagged UbiquE24C/A28C, displaying a lower average signal-to-noise ratio and a minor species at approximately 30–40% population (Figure S7 in the Supporting Information; PCSs for the major species are presented in Table S3 in the Supporting Information). Moreover, although the calculated metal position for the major species (Figure S8 in the Supporting Information) was indeed found to be consistent with additional coordination by E24 for this tagged mutant (as opposed to D32 in the case of the lanthanides^[12]), the corresponding $\Delta\chi$ tensor [$\chi_{Ax} = (-4.5 \pm 0.6) \times 10^{-32} \text{ m}^3$, $\chi_{Rh} = (-2.4 \pm 0.4) \times 10^{-32} \text{ m}^3$] is still smaller than that for the Co^{II} -bis-IDA motif.

Based on the magnitude of the $\Delta\chi$ tensor for the Co^{II} -bis-IDA motif, nuclei lying up to 40 Å from the Co^{II} centre along the largest principal (Z) axis of the $\Delta\chi$ tensor can be interrogated by a PCS of 0.08 ppm. The sharp lines and small PRE effects observed in the spectrum of the Co^{II} -bis-IDA tagged ubiquitin permitted high signal-to-noise spectra for almost all observed spins to be recorded in less than half-an-hour. This is considerably faster than the case for the stronger lanthanides, such as the Dy^{III} -loaded single-IDA-tagged ubiquitin,^[12] owing to the excess of free Dy^{III} ions required to account for weaker metal ion binding, a much larger PRE effect compared to Co^{II} , and greater variability in observed linewidths. The highly rigid Co^{II} -bis-IDA motif is therefore applicable to moderately sized proteins and their complexes, even though the inherent paramagnetic strength of high-spin Co^{II} is rather small compared to that of some of the lanthanide ions. For comparison, the magnitude of the $\Delta\chi$ tensor is similar to that observed for the Yb^{III} -loaded two-arm “CLaNP-5” LBT^[20] attached to several sites on the FKBP12 protein.^[21] Notably, the bis-IDA tag ensures a fully saturated coordination motif that reduces the likelihood of Co^{II} -mediated protein dimerisation.

The addition of a “small” lanthanide ion (Tm^{III}) to the bis-IDA tagged ubiquitin produced many sizeable PCSs, consistent with a large $\Delta\chi$ tensor (Figure S9 in the Supporting Information) but multiple species were evident, resulting in complex spectra that were unsuitable for PCS analysis. Several of the observed PCSs were consistent with preferential binding to a single IDA attached to A28C,^[12] with additional coordination from a nearby aspartic acid residue (D32), although the com-

plex and heterogeneous spectra indicated that a mixture of coordination species were present. This observation is likely due to a combination of the presence of a number of low denticity metal ion-binding groups, and the preference of lanthanide ions for high coordination numbers, allowing for multiple binding possibilities with comparable affinity. Addition of a larger lanthanide ion (Tb^{III}) similarly gave rise to multiple PCSs (Figure S9). Both spectra were also noticeably broader and of lower quality compared to the high-quality spectra observed previously for the A28C-IDA/D32 motif.^[12] These observations suggest that lanthanide ions cannot generally be bound by the bis-IDA motif in a manner suitable for PCS-based structural studies. The smaller ionic radius of Co^{II} (74 pm) likely matches the dimensions of the bis-IDA “cage” better than the lanthanide ions, the radii of which vary from 101 pm for Ce^{III} to 87 pm for Yb^{III} .^[22]

Compared to the single IDA-lanthanide ion tagging strategy^[12] the analysis of the data is perhaps more onerous, owing to the fact that: 1) assignment of the spectrum is not aided by the scope to record spectra for several different metal ions of varying paramagnetic strength and similarly oriented $\Delta\chi$ tensors, and 2) there is an increase in peak number upon adding the diamagnetic Zn^{II} reference ion to the protein in substoichiometric amounts. Nevertheless, in the absence of a suitable structural model, PCS assignments can be potentially made directly by performing standard triple resonance experiments on double-labelled samples, or aided by acquiring spectra at various temperatures.^[23] The capacity to define the $\Delta\chi$ tensor from just PCSs of amides of the helix to which the Co^{II} -bis-IDA motif is attached is a unique feature that will facilitate structural analyses and de novo structure determination.

In conclusion, attachment of two copies of the IDA-based tag, **1**, in an $i, i+4$ configuration on a protein α -helix results in an autonomous, chirally pure, hexadentate metal ion-binding motif suitable for complexation of Co^{II} (and potentially other octahedral paramagnetic ions such as Ni^{II} and Mn^{II}). The tagging strategy produces measurable PCSs for nuclei located up to 40 Å or more from the Co^{II} centre. The small PRE effect of Co^{II} , combined with the high rigidity of the Co^{II} -bis-IDA motif, also means that very large, structurally accurate PCSs can be observed for nuclei positioned close to the metal centre (≈ 8 Å away). A Co^{II} -bis-IDA-based tagging strategy will thus be able to interrogate spin coordinates for nuclei located over quite a broad range of distances. Compared to previous approaches, the new tagging strategy does not rely on the Co^{II} making additional contacts with nearby side-chains to impart rigidity and therefore represents a robust and predictable general method for generating practically useful PCS data.

Experimental Section

Preparation of compound **1**

To a solution of (*R*)-2-((carboxymethyl)amino)-3-mercaptopropanoic acid^[12] (15 mg, 0.084 mmol) in a 1:1 (v/v) mixture of MeOH/H₂O (5 mL) was added 2,2'-dithiodipyridine (40 mg, 0.18 mmol) and the resulting mixture stirred for 2 h at room temperature. After this

time, the crude was purified by preparative RP-HPLC (0.1% TFA, H₂O/CH₃CN gradient) (t_{R} = 14.42 min) and the combined fractions lyophilised to obtain a clear oil (24 mg, 89%). ¹H NMR (CD₃OD, 400 MHz): δ = 3.51–3.67 (m, 2H, C _{β} H₂), 4.12 (dd, 2H, J = 4.4/16.8 Hz, CH₂), 4.39 (t, 1H, J = 6.8 Hz, C _{α} H), 7.36 (dd, 1H, J = 4.8/6.6 Hz, CH), 7.59 (d, 1H, J = 8.0 Hz, CH), 7.80 (ddd, 1H, J = 1.6/8.0/7.8 Hz, CH), 8.62 ppm (d, 1H, J = 4.4 Hz, CH); ¹³C NMR (CD₃OD, 101 MHz): δ = 38.15 (C _{β} H₂), 46.31 (CH₂), 58.76 (C _{α} H), 123.73 (CH), 123.94 (CH), 139.43 (CH), 151.07 (CH), 158.94 (C), 169.44 (CO), 169.60 ppm (CO); ESI-MS (m/z): 289.00 [M +H]⁺ (100%); analytical RP-HPLC (214 nm): t_{R} = 3.86 min, > 99% purity.

Tagging of ubiquitin E24C/A28C with **1**

Uniformly ¹⁵N-labelled samples of the E24C/A28C mutant of human ubiquitin were prepared as previously described.^[12] Ligation with **1** was achieved by the addition of a tenfold excess of **1** to a 30 μM sample of ubiquitin E24C/A28C dissolved in 50 mM HEPES (pH 7.5) followed by stirring at room temperature for 1 h. Excess tag was removed by passage over a PD-10 column equilibrated with 50 mM HEPES (pH 7.5).

Protein NMR spectroscopy

All protein NMR spectra were acquired on either a Varian Inova or Bruker Avance 600 MHz NMR spectrometer equipped with a cryoprobe and a single Z axis gradient. Titrations were performed by addition of aliquots from 5 mM stock solutions of ZnCl₂, CoCl₂ or the triflate salts of selected lanthanides. Final protein concentration was typically 70–100 μM dissolved in a HEPES buffer (50 mM) at pH 7.5 containing 10% D₂O. All spectra were performed in 5 mm OD Shigemi microcell NMR tubes (Shigemi Inc.). The 2D ¹⁵N-fast-HSQC were acquired typically with a $t_{1\text{max}}$ (¹⁵N) of 50–62 ms and a $t_{2\text{max}}$ (¹H) of 142 ms. For assignments of the Zn^{II}-loaded sample, a 3D ¹⁵N-separated NOESY-Fast-HSQC spectrum was recorded on a 200 μM sample using a mixing time of 170 ms.

Modelling of Co^{II}-loaded ubiquitin E24C/A28C

Modelling of the bis-IDA cage was performed using the Schrödinger Suite 2014 (www.schrodinger.com) through the Maestro interface (Maestro, version 9.7, Schrödinger, LLC, New York, 2014) and with XPLOR-NIH^[24] (see details in Supporting Information). Figures were generated using MOLMOL^[25] or UCSF Chimera.^[26]

Acknowledgements

B.G. would like to thank the Australian Research Council (Future Fellowship FT130100838 and Discovery Grants DP120100561 and DP150100383). We thank Prof. Gottfried Otting for reading a preliminary version of the manuscript and for some helpful suggestions.

Keywords: alpha-helix • cobalt • paramagnetic NMR spectroscopy • protein structures • pseudo-contact shifts

- [1] a) I. Bertini, C. Luchinat, G. Parigi, *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 249–273; b) J. Koehler, J. Meiler, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 360–389; c) T. Ikegami, L. Verdier, P. Sakhaei, S. Grimme, B. Pescatore, K. Saxena, K. M. Fiebig, C. Griesinger, *J. Biomol. NMR* **2004**, *29*, 339–349.

- [2] a) I. Baig, I. Bertini, C. D. Bianco, Y. K. Gupta, *Biochemistry* **2004**, *43*, 5562–5573; b) G. Pintacuda, A. Y. Park, M. A. Keniry, N. E. Dixon, G. Otting, *J. Am. Chem. Soc.* **2006**, *128*, 3696–3702; c) G. Pintacuda, M. John, X. C. Su, G. Otting, *Acc. Chem. Res.* **2007**, *40*, 206–212; d) Y. Hiruma, M. S. Hass, Y. Kikui, W. M. Liu, B. Olmez, S. P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Löhr, H. Schwalbe, M. Nojiri, M. Ubbink, *J. Mol. Biol.* **2013**, *425*, 4353–4365.
- [3] a) I. Bertini, C. Del Bianco, I. Gelis, N. Katsaros, C. Luchinat, G. Parigi, M. Peana, A. Provenzano, M. A. Zoroddu, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6841–6846; b) X. Xu, P. H. Keizers, W. Reinie, F. Hannemann, R. Bernhardt, M. Ubbink, *J. Biomol. NMR* **2009**, *43*, 247–254.
- [4] a) H. Yagi, K. B. Pilla, A. Maleckis, B. Graham, T. Huber, G. Otting, *Structure* **2013**, *21*, 883–890; b) C. Schmitz, R. Vernon, G. Otting, D. Baker, T. Huber, *J. Mol. Biol.* **2012**, *416*, 668–677.
- [5] D. J. Crick, J. X. Wang, B. Graham, J. D. Swarbrick, H. R. Mott, D. Nietlispach, *J. Biomol. NMR* **2015**, *61*, 197–207.
- [6] a) M. Nitz, K. J. Franz, R. L. Maglathlin, B. Imperiali, *ChemBioChem* **2003**, *4*, 272–276; b) X. C. Su, K. McAndrew, *J. Am. Chem. Soc.* **2008**, *130*, 1681–1687; c) T. Saio, K. Ogura, M. Yokochi, Y. Kobashigawa, F. Inagaki, *J. Biomol. NMR* **2009**, *44*, 157–166; d) K. Barthelmes, A. M. Reynolds, E. Peisach, H. R. Jonker, N. J. DeNunzio, K. N. Allen, B. Imperiali, H. Schwalbe, *J. Am. Chem. Soc.* **2011**, *133*, 808–819.
- [7] a) W. M. Liu, M. Overhand, M. Ubbink, *Coord. Chem. Rev.* **2014**, *273*, 2–12; b) X. C. Su, G. Otting, *J. Biomol. NMR* **2010**, *46*, 101–112.
- [8] a) A. Donaire, J. Salgado, J. M. Moratal, *Biochemistry* **1998**, *37*, 8659–8673; b) C. Dennison, K. Sato, *Inorg. Chem.* **2004**, *43*, 1502–1510.
- [9] M. Gochin, *J. Biomol. NMR* **1998**, *12*, 243–257.
- [10] B. Man, X. C. Su, H. Liang, S. Simonsen, T. Huber, B. Messerle, G. Otting, *Chem. Eur. J.* **2010**, *16*, 3827–3832.
- [11] T. H. D. Nguyen, K. Ozawa, M. Stanton-Cook, R. Barrow, T. Huber, G. Otting, *Angew. Chem. Int. Ed.* **2011**, *50*, 692–694; *Angew. Chem.* **2011**, *123*, 718–720.
- [12] J. D. Swarbrick, P. Ung, S. Chhabra, B. Graham, *Angew. Chem. Int. Ed.* **2011**, *50*, 4403–4406; *Angew. Chem.* **2011**, *123*, 4495–4498.
- [13] J. D. Swarbrick, P. Ung, X. C. Su, A. Maleckis, S. Chhabra, T. Huber, B. Graham, *Chem. Commun.* **2011**, *47*, 7368–7370.
- [14] H. Yagi, A. Maleckis, G. Otting, *J. Biomol. NMR* **2013**, *55*, 157–166.
- [15] a) Q. Z. Zhang, C. Z. Lu, W. B. Yang, *J. Coord. Chem.* **2006**, *59*, 837–844; b) D. Tarn, M. Xue, J. I. Zink, *Inorg. Chem.* **2013**, *52*, 2044–2049.
- [16] C. Schmitz, M. J. Stanton-Cook, X. C. Su, G. Otting, T. Huber, *J. Biomol. NMR* **2008**, *41*, 179–189.
- [17] A. S. Maltsev, A. Grishaev, J. Roche, M. Zasloff, A. Bax, *J. Am. Chem. Soc.* **2014**, *136*, 3752–3755.
- [18] D. Shishmarev, G. Otting, *J. Biomol. NMR* **2013**, *56*, 203–216.
- [19] I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, *ChemBioChem* **2005**, *6*, 1536–1549.
- [20] P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand, M. Ubbink, *J. Am. Chem. Soc.* **2008**, *130*, 14802–14812.
- [21] J. Y. Guan, P. H. J. Keizers, W.-M. Liu, F. Loehr, S. P. Skinner, E. A. Heeneman, H. Schwalbe, M. Ubbink, G. D. Siegal, *J. Am. Chem. Soc.* **2013**, *135*, 5859–5868.
- [22] R. D. Shannon, *Acta. Crystallogr. A* **1976**, *32*, 751–767.
- [23] M. Allegrozzi, I. Bertini, M. B. L. Janik, Y. M. Lee, G. Liu, C. Luchinat, *J. Am. Chem. Soc.* **2000**, *122*, 4154–4161.
- [24] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65–73.
- [25] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51–55.
- [26] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

Received: August 9, 2015

Revised: December 2, 2015

Published online on December 17, 2015



Cite this: *Chem. Commun.*, 2016, 52, 7954

Received 17th March 2016,
Accepted 26th May 2016

DOI: 10.1039/c6cc02325h

www.rsc.org/chemcomm

Enantiomeric two-armed lanthanide-binding tags for complementary effects in paramagnetic NMR spectroscopy†

Michael D. Lee,^a Matthew L. Dennis,^{ab} James D. Swarbrick^{*a} and Bim Graham^{*a}

Two-armed lanthanide-binding tags induce significant, long-range paramagnetic effects in the NMR spectra of attached proteins. An enantiomeric pair of rigid, two-armed, cyclen-based tags are reported that produce markedly different effects from the same tagging site, allowing for the measurement of complementary paramagnetic restraints for structural studies.

Paramagnetic lanthanide ions induce unique effects in the NMR spectra of bound proteins, including paramagnetic relaxation enhancement (PRE), residual dipolar couplings (RDCs) and pseudo-contact shifts (PCSs).^{1–3} These can be measured to yield valuable long-range restraints that aid characterisation of protein structure^{4,5} and dynamics^{6,7} as well as protein–protein^{8,9} and protein–ligand interactions.^{10–13}

Over the past decade, lanthanide-binding tags (LBTs) have been developed for the site-specific attachment of lanthanide ions (Ln³⁺) to proteins in order to utilise these paramagnetic restraints for structural investigations.^{1,14} One of the key aspects to a LBT's performance is the rigidity of the lanthanide ion relative to the protein, as excessive movement of the lanthanide will cause averaging of the paramagnetic effects, reducing their utility. The attachment of a lanthanide to a protein by two points of conjugation, using “two-armed” LBTs, has been shown to be an excellent method of rigidifying the metal ion position relative to the protein and generally produces reliably

larger paramagnetic effects than single-point-attachment (single-armed) LBTs.^{15–19}

Structural studies using PCSs can benefit from the attachment of a LBT at different tagging sites.^{20,21} This provides a greater PCS “coverage” over the protein and gives rise to structurally distinct PCSs for individual nuclei, which can reduce structural ambiguities associated with the shape and symmetry of a single magnetic susceptibility anisotropy ($\Delta\chi$ -) tensor. Another useful approach to resolve such ambiguities is to attach different LBTs to the same tagging site of a protein.^{22,23} As long as the tags differ sufficiently from one another in terms of the orientation of their $\Delta\chi$ - and alignment tensors, structurally distinct PCSs and RDCs can be obtained, providing a greater number of paramagnetic restraints per tagging site and reducing the number of protein mutants required for structural studies. This may be particularly advantageous when utilising two-armed LBTs, as the generation of viable double-cysteine mutant proteins can be challenging, due to the unfavourable effects that proximal cysteine residue can have on protein expression, folding and stability.^{21,24} However, to date, this approach has scarcely been explored using two-armed LBTs, which may reflect in part the fact that all existing two-armed tags are based upon a similar achiral design.

We now report a pair of compact, hydrophilic, enantiomeric two-armed LBTs, **T1** and **T2** (Fig. 1), that reliably produce large, yet substantially different $\Delta\chi$ - and alignment tensors when bound to the same protein mutant, as demonstrated by attachment of the tags to two “test” proteins, human ubiquitin and *S. aureus* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK). The tags are two-armed variants of the previously reported cyclen-based tag, **C5**,²⁵ featuring pairs of pyridyl disulphide-activated linkers for conjugation to two cysteine residues *via* disulphide bond formation (see ESI† for synthetic details). Due to their enantiomeric nature, the tags interact differently with the local chiral protein environment, giving rise to structurally distinct PCSs and RDCs, and thereby providing a rich source of complementary distance and angular information from a single tagging site.

The Tm³⁺, Yb³⁺ and Y³⁺-loaded forms of **T1** and **T2** were first attached to the α -helix of a ubiquitin E24C/A28C double-cysteine

^a Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia.

^b CSIRO Biosciences Program, Parkville, VIC 3052, Australia

† Electronic supplementary information (ESI) available: Synthesis of **T1** and its lanthanide complexes, HRMS of **T2**-Yb³⁺, 1D ¹H and ¹³C NMR spectra of novel compounds, details for expression, purification and tagging of ubiquitin E24C/A28C and HPPK K76C/C80, details for protein NMR spectroscopy, $\Delta\chi$ -tensor and alignment tensor determination, ¹⁵N-HSQC spectra of differently tagged ubiquitin E24C/A28C and HPPK K76C/C80, full tables of $\Delta\chi$ -tensor and alignment tensor properties, correlations between measured and calculated PCSs and RDCs, correlations between PCSs or RDCs measured with **T1** and **T2**, Sanson-Flamsteed projections of $\Delta\chi$ -tensor and alignment tensor error analyses, structures of ubiquitin and HPPK with determined $\Delta\chi$ -tensor axes, correlations between PCSs measured at pH 8.0 and 6.5 for **T1**- and **T2**-tagged ubiquitin. See DOI: 10.1039/c6cc02325h

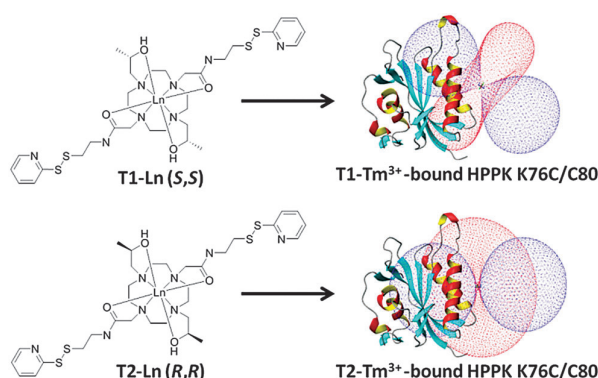


Fig. 1 The structures of **T1-Ln** and **T2-Ln** and isosurface representations of the $\Delta\chi$ -tensors of their Tm³⁺ complexes bound to the same HPPK K76C/C80 mutant protein.

mutant (see ESI† for details). Only a single set of PCSs are evident in the spectra of the Tm³⁺- and Yb³⁺-tagged protein relative to the diamagnetic Y³⁺-tagged samples (Fig. 2). For both tags, the majority of the observed PCSs are positive, yet the PCSs for specific nuclei are different with each enantiomer. The $\Delta\chi$ -tensor properties were determined by fitting the measured PCSs (Table S5, ESI†) to the structure of ubiquitin (PDB 2MJB²⁶) using Numbat,²⁷ revealing very sizable $\Delta\chi$ -tensors with low Q-values (Fig. S9, ESI† and Table 1).

For the **T1-Tm³⁺** complex, the $\Delta\chi_{ax}$ component is more than twice the size of that of **C5-Tm³⁺**-tagged ubiquitin A28C under the same sample conditions,²⁵ demonstrating the significant effect that two-point attachment has on the apparent $\Delta\chi$ -tensor magnitude, although the change from a hydroxyl to an amide

Table 1 $\Delta\chi$ -tensor properties of protein-bound **T1/T2** tags at pH 8.0^a

Protein	Tag	Ln ³⁺	#PCS	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	Q
Ubi	T1	Tm ³⁺	32	49.9 (0.4)	8.8 (0.5)	0.02
		Yb ³⁺	52	−9.1 (0.2)	−4.2 (0.1)	0.05
	T2	Tm ³⁺	41	35.3 (0.5)	4.0 (0.4)	0.03
		Yb ³⁺	43	5.5 (0.1)	2.1 (0.2)	0.07
HPPK	T1	Tm ³⁺	67	41.0 (1.0)	6.8 (0.7)	0.08
		Yb ³⁺	117	7.1	4.6	0.07
	T2	Tm ³⁺	94	58.0 (0.9)	9.9 (1.4)	0.05
		Yb ³⁺	106	9.5	5.4	0.03

^a The axial and rhombic components of the $\Delta\chi$ -tensors are reported in units of 10^{−32} m³. Standard deviations (in brackets) were determined from random removal of 10% of the PCSs and recalculating the $\Delta\chi$ -tensors 1000 times, in some cases the z and y axes of the tensors were of similar magnitude and swapped in different fits, thus standard deviations were not determined. See Table S2 (ESI†) for metal ion coordinates, Euler angles and ubiquitin tensors at pH 6.5.

metal-coordinating group will also affect the tags' paramagnetic properties. This particular $\Delta\chi$ -tensor is sufficiently large that a nuclei located 50 Å away from the Tm³⁺ ion along the Z axis of the $\Delta\chi$ -tensor would have a predicted PCS of 0.2 ppm. Notably, the increased magnitude of the $\Delta\chi$ -tensor is less pronounced for the Yb³⁺ complex, as the $\Delta\chi_{ax}$ of **T1-Yb³⁺**-tagged ubiquitin E24C/A28C is only 36% larger than that of **C5-Yb³⁺**-tagged ubiquitin A28C. In practice, we found that the assignment of the very large Tm³⁺ PCSs was greatly assisted by assignment and determination of the Yb³⁺ PCSs and $\Delta\chi$ -tensor properties, providing an approximate metal ion coordinate to facilitate the Tm³⁺ $\Delta\chi$ -tensor determination.

The **T1** and **T2** complexes were also attached to *S. aureus* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase²⁸ K76C/C80 to further assess the tags' performance on a second protein helix. Strikingly, **T1-Tm³⁺**-tagged HPPK yielded relatively small PCSs for such a strongly paramagnetic metal attached by a two-armed tag to a regular α -helix (Fig. S9, S10 and Table S6, ESI†). However, subsequent analysis of the PCSs revealed large $\Delta\chi$ -tensors for the complexes of both **T1**- and **T2**-tagged HPPK (Table 1).

To assess the tags' rigidity, ¹D_{HN} RDCs up to 16.1/16.7 and 24.8/31.8 Hz were measured at 600 MHz for the **T1/T2-Tm³⁺**-tagged ubiquitin and HPPK, respectively. The alignment tensors were determined by fitting the measured RDCs (Tables S7 and S8, ESI†) to the structures of ubiquitin and HPPK, respectively, with good-to-fair Q-values (0.07–0.22) (Fig. S11 and Table S3, ESI†). The axial components of the alignment tensors are, on average, 88% of the magnitude of the PCS-determined $\Delta\chi_{ax}$ components, supportive of the tags' rigid attachment to each protein. The smaller tensor sizes and larger Q-values obtained from the fitting of the RDC *versus* the PCS data are associated with the greater sensitivity of RDCs to minor differences in protein structure and residual tag movements, as has been noted previously.^{29,30} For both the PCS- and RDC-derived $\Delta\chi$ - and alignment tensors, the axial terms are typically more well-defined than the rhombic ones, due to the dominant axial character and orientation of the tensors (Fig. S12, ESI† and Table 1).

Comparison of the PCSs or RDCs measured with **T1**- and **T2**-tagged ubiquitin and HPPK are shown in Fig. 3 and Table S4 (ESI†). The PCSs of each enantiomer bound to the same protein

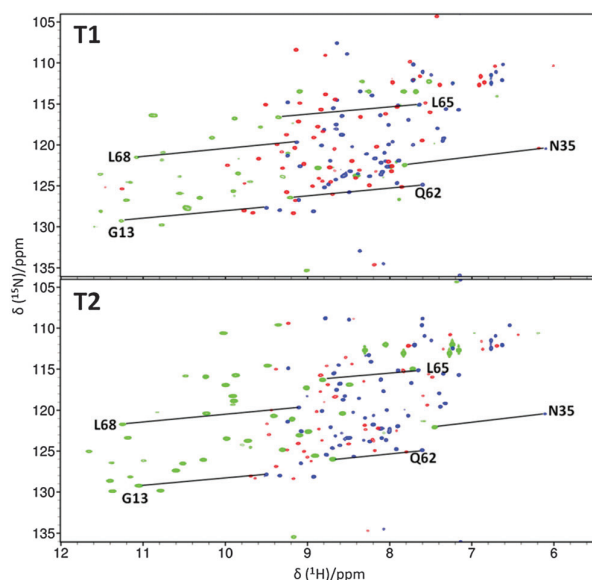


Fig. 2 Overlays of ¹⁵N-HSQC spectra of **T1**- (top) and **T2**- (bottom) tagged ubiquitin E24C/A28C, loaded with either Y³⁺ (blue), Tm³⁺ (green) or Yb³⁺ (red). The spectra were recorded at 25 °C and pH 8.0 at a ¹H NMR frequency of 600 MHz. Selected PCSs are indicated with solid lines.

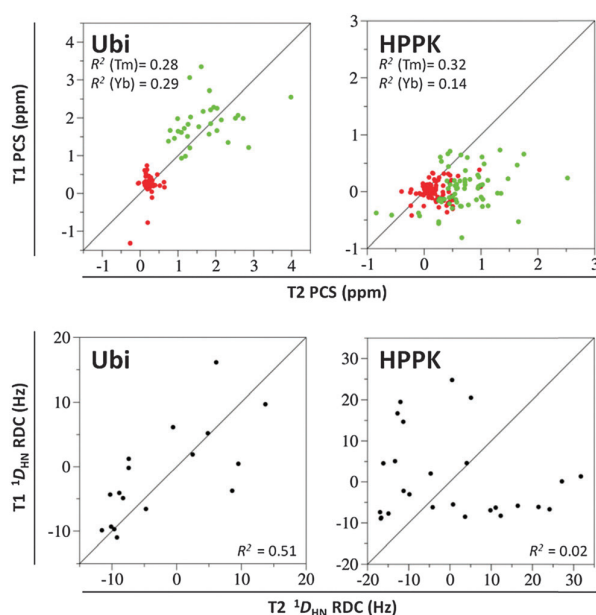


Fig. 3 Correlations between PCSs or RDCs measured with **T1** and **T2** tags at pH 8.0, loaded with either Tm³⁺ (green and black) or Yb³⁺ (red). Only PCSs or RDCs that were measured with both tags are shown. Solid lines represent perfect correlation.

show very poor correlation with one another (R^2 ranging from 0.14–0.32). Surprisingly, the RDCs of the **T1/T2-Tm**³⁺-tagged ubiquitin showed a greater (yet still poor) correlation than the respective PCSs (R^2 of 0.51 and 0.28 for the RDCs and PCSs, respectively). However, this may be due to the fewer number of residues for which an RDC could be measured with both tags, compared to the case for the PCSs (17 RDCs *cf.* 31 PCSs). In contrast, the RDCs measured for **T1/T2-Tm**³⁺-tagged HPPK show extremely poor correlation (R^2 of 0.02), due to the different orientations of the alignment tensors of the tagged HPPK (*vide infra*). Thus, the tags clearly provide distinct and complementary structural information when attached to the same double-cysteine mutant protein.

The PCS-determined metal ion positions for each tag are quite similar (only 1.6 and 2.2 Å apart for ubiquitin and HPPK, respectively) and lie between the two proximal cysteine residues for both proteins. The *Z* axes of the **T1-Tm**³⁺ and **T2-Tm**³⁺ $\Delta\chi$ -tensors are tilted by 18° and 54° relative to one another for ubiquitin and HPPK, respectively (Fig. S12 and S13, ESI†). For **T1**- and **T2**-tagged ubiquitin and **T2**-tagged HPPK, the *Z* axes of the $\Delta\chi$ -tensors are orientated roughly perpendicular to the attached protein α -helix, resulting in the majority of the protein falling within a single lobe of the $\Delta\chi$ -tensor and producing predominantly positive PCSs. For **T1**-tagged HPPK, the *Z* axis of the $\Delta\chi$ -tensor is noticeably tilted, such that the protein lies between the different tensor lobes and both positive and negative PCSs are observed.

Interestingly, the **T1** tag's $\Delta\chi$ -tensor magnitudes and orientations are similar for both proteins, while the **T2** tag's properties vary significantly between ubiquitin and HPPK. For one or both of

the proteins, the **T2** tag may be forming more significant secondary interactions with the local protein environment, influencing the tag's properties. The work of Sessoli and co-workers has demonstrated that even subtle changes in coordination environment (*e.g.* re-positioning of the two hydrogens of a coordinated water molecule) can potentially have a significant impact on the magnetic anisotropy of a lanthanide complex.³¹ The sensitivity of the tags' properties to the protein environment indicates that they may not be suitable for the *de novo* prediction of PCSs, as has been done using the achiral **CLaNP-5** tag.^{16,21}

The **T1/T2** tags feature coordinating hydroxyl and amide pendants (and potentially an aqua ligand) that may exist in protonated (neutral) and deprotonated forms,^{32,33} and therefore their magnetic properties (including tensor size and orientation) could be sensitive to pH. To explore this possibility, spectra of the **T1/T2**-tagged ubiquitin samples were rerecorded at pH 6.5 (Fig. S14, ESI†). The magnitude of the observed PCSs and determined tensors are on average 40% smaller than those obtained at pH 8.0 (Fig. S15 and Table S2, ESI†). The metal ion positions and tensor orientations are not significantly changed, suggesting that recording spectra at more than one pH may be beneficial for assignment of the large **T1/T2-Tm**³⁺ PCSs (Fig. S16, ESI†). Importantly, the tags produced only a single set of PCSs at both pH 6.5 and 8, allowing for straightforward analysis.

Few two-armed LBTs have previously been reported, most notably the **CLaNP-5**, -7, and -9 tags.^{15,16,34,35} Thus, the use of multiple two-armed tags at each attachment site is already possible, but to our knowledge has not been taken advantage of in any structural studies. When bound to a Paz E51C/E54C mutant protein, the Tm³⁺ complexes of **CLaNP-5** and **CLaNP-7** produced PCSs with an R^2 of 0.84 relative to each other,³⁴ significantly higher than that observed for **T1** and **T2** when attached to ubiquitin or HPPK. The **CLaNP-7** tag displays a pH dependence that can produce differently orientated $\Delta\chi$ tensors when a histidine residue is located near the tagging site.³⁴ However, this requires more detailed prior structural knowledge of the protein to be tagged and for the protein to be stable across a varied pH range. The parent (**S**)-**THP-Yb**³⁺ complex of **T1** has been reported to adopt a twisted square anti-prismatic coordination geometry in solution,³² in contrast to the square anti-prismatic geometry of the chelates upon which the **CLaNP** tags are based.^{36–38} Thus, it is likely that the different tags would also produce complementary information from the same site, which may be desirable for situations where very few (or only a single) double-cysteine mutant proteins can be produced. Although the effective “linkers” between the protein-bound **T1/T2** are the same as those of the **CLaNP-5** and **CLaNP-7** tags, the chiral elements of the **T1/T2** design may help shed light on the source of the residual motions of the **CLaNP-5** tag, recently reported in relaxation dispersion NMR experiments.³⁹

In conclusion, we have demonstrated that the enantiomeric **T1/T2** tags, which can be readily synthesised in parallel, can provide long-range complementary structural information from the same tagging site, thus requiring fewer double-cysteine protein mutants for structural investigations. This reflects the fact that attachment of the tags produces a pair of diastereomeric

protein-tag conjugates. Although the tags' paramagnetic effects vary with factors such as the sample pH and local protein environment, their rigid two-point attachment produces reliably large magnetic susceptibility anisotropy and alignment tensors from which PCSs and RDCs can be measured. Thus, the tags represent attractive new tools for protein structural studies. In addition to studies of ligand binding events and associated structural changes, they may prove useful for rigid body docking of large protein–protein complexes.

An Australian Research Council Future Fellowship (FT1301100838) to B. G. is gratefully acknowledged.

Notes and references

- W.-M. Liu, M. Overhand and M. Ubbink, *Coord. Chem. Rev.*, 2014, **273–274**, 2–12.
- G. Otting, *Annu. Rev. Biophys.*, 2010, **39**, 387–405.
- I. Bertini, C. Luchinat and G. Parigi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2002, **40**, 249–273.
- K. B. Pilla, G. Otting and T. Huber, *J. Mol. Biol.*, 2016, **428**, 522–532.
- D. J. Crick, J. X. Wang, B. Graham, J. D. Swarbrick, H. R. Mott and D. Nietlispach, *J. Biomol. NMR*, 2015, **61**, 197–207.
- W. Andraóje, C. Luchinat, G. Parigi and E. Ravera, *J. Phys. Chem. B*, 2014, **118**, 10576–10587.
- L. Russo, M. Maestre-Martinez, S. Wolff, S. Becker and C. Griesinger, *J. Am. Chem. Soc.*, 2013, **135**, 17111–17120.
- Y. Hiruma, M. A. S. Hass, Y. Kikui, W.-M. Liu, B. Olmez, S. P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Löhr, H. Schwalbe, M. Nijiri and M. Ubbink, *J. Mol. Biol.*, 2013, **425**, 4353–4365.
- T. Saio, M. Yokochi, H. Kumeta and F. Inagaki, *J. Biomol. NMR*, 2010, **46**, 271–280.
- S. P. Skinner, W.-M. Liu, Y. Hiruma, M. Timmer, A. Blok, M. A. S. Hass and M. Ubbink, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 9022–9027.
- T. Zhuang, H. Lee, B. Imperiali and J. H. Prestegard, *Protein Sci.*, 2008, 1220–1231.
- T. Saio, K. Ogura, H. Kumeta, Y. Kobashigawa, K. Shimizu, M. Yokochi, K. Kodama, H. Yamaguchi, H. Tsujishita and F. Inagaki, *Sci. Rep.*, 2015, **5**, 16685.
- Á. Canales, Á. Mallagaray, M. Á. Berbis, A. Navarro-Vázquez, G. Domínguez, F. J. Cañada, S. André, H.-J. Gabius, J. Pérez-Castells and J. Jiménez-Barbero, *J. Am. Chem. Soc.*, 2014, **136**, 8011–8017.
- J. Koehler and J. Meiler, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **59**, 360–389.
- P. H. J. Keizers, J. F. Desreux, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2007, **129**, 9292–9303.
- P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**, 14802–14812.
- T. Saio, K. Ogura, M. Yokochi, Y. Kobashigawa and F. Inagaki, *J. Biomol. NMR*, 2009, **44**, 157–166.
- J. D. Swarbrick, P. Ung, X. Su, A. Maleckis, S. Chhabra, T. Huber, G. Otting and B. Graham, *Chem. Commun.*, 2011, **47**, 7368–7370.
- J. D. Swarbrick, P. Ung, M. L. Dennis, M. D. Lee, S. Chhabra and B. Graham, *Chem. – Eur. J.*, 2016, **22**, 1228–1232.
- H. Yagi, K. B. Pilla, A. Maleckis, B. Graham, T. Huber and G. Otting, *Structure*, 2013, **21**, 883–890.
- J.-Y. Guan, P. H. J. Keizers, W.-M. Liu, F. Loehr, S. P. Skinner, E. A. Heeneman, H. Schwalbe, M. Ubbink and G. D. Siegal, *J. Am. Chem. Soc.*, 2013, **135**, 5859–5868.
- L. De La Cruz, T. H. D. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2011, **133**, 19205–19215.
- Y. Kobashigawa, T. Saio, M. Ushio, M. Sekiguchi, M. Yokochi, K. Ogura and F. Inagaki, *J. Biomol. NMR*, 2012, **53**, 53–63.
- A. R. Camacho-Zarco, F. Munari, M. Wegstroth, W.-M. Liu, M. Ubbink, S. Becker and M. Zweckstetter, *Angew. Chem., Int. Ed.*, 2015, **54**, 336–339.
- M. D. Lee, C. Loh, J. Shin, S. Chhabra, M. L. Dennis, G. Otting, J. D. Swarbrick and B. Graham, *Chem. Sci.*, 2015, **6**, 2614–2624.
- A. S. Maltsev, A. Grishaev, J. Roche, M. Zasloff and A. Bax, *J. Am. Chem. Soc.*, 2014, **136**, 3752–3755.
- C. Schmitz, M. J. Stanton-Cook, X.-C. Su, G. Otting and T. Huber, *J. Biomol. NMR*, 2008, **41**, 179–189.
- S. Chhabra, O. Dolezal, B. M. Collins, J. Newman, J. S. Simpson, I. G. Macreadie, R. Fernley, T. S. Peat and J. D. Swarbrick, *PLoS One*, 2012, **7**, e29444.
- Y. Yang, J. Wang, Y. Pei and X. Su, *Chem. Commun.*, 2015, **51**, 2824–2827.
- F. Peters, M. Maestre-Martinez, A. Leonov, L. Kovačič, S. Becker, R. Boelens and C. Griesinger, *J. Biomol. NMR*, 2011, **51**, 329–337.
- G. Cucinotta, M. Perfetti, J. Luzon, M. Etienne, P.-E. Car, A. Caneschi, G. Calvez, K. Bernot and R. Sessoli, *Angew. Chem., Int. Ed.*, 2012, **51**, 1606–1610.
- M. Lelli, G. Pintacuda, A. Cuzzola and L. Di Bari, *Chirality*, 2005, **17**, 201–211.
- K. Chin and J. Morrow, *Inorg. Chem.*, 1994, **33**, 5036–5041.
- W.-M. Liu, P. Keizers, M. Hass, A. Blok, M. Timmer, A. J. Sarris, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2012, **132**, 17306–17313.
- W.-M. Liu, S. P. Skinner, M. Timmer, A. Blok, M. A. S. Hass, D. V. Filippov, M. Overhand and M. Ubbink, *Chem. – Eur. J.*, 2014, **20**, 6256–6258.
- M. Polášek, J. Kotek, P. Hermann, I. Čisarová, K. Binnemans and I. Lukeš, *Inorg. Chem.*, 2009, **48**, 466–475.
- M. Polášek, J. Rudovský, P. Hermann, I. Lukes, L. Vander Elst and R. N. Muller, *Chem. Commun.*, 2004, 2602–2603.
- M. Woods, G. E. Kiefer, S. Bott, A. Castillo-Muzquiz, C. Eshelbrenner, L. Michaudet, K. McMillan, S. D. K. Mudigunda, D. Ogrin, G. Tircsó, S. Zhang, P. Zhao and A. D. Sherry, *J. Am. Chem. Soc.*, 2004, **126**, 9248–9256.
- M. A. S. Hass, W.-M. Liu, R. V. Agafonov, R. Otten, L. A. Phung, J. T. Schilder, D. Kern and M. Ubbink, *J. Biomol. NMR*, 2015, **61**, 123–136.

This Agreement between Matthew L Dennis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	4006310621344
License date	Dec 12, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Drug Discovery
Licensed Content Title	The influence of the 'organizational factor' on compound quality in drug discovery
Licensed Content Author	Paul D. Leeson and Stephen A. St-Gallay
Licensed Content Date	Oct 1, 2011
Licensed Content Volume	10
Licensed Content Issue	10
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 7
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Inhibitors of the folate pathway enzymes HPPK and DHPS
Expected completion date	Dec 2016
Estimated size (number of pages)	191
Requestor Location	Matthew L Dennis 381 Royal Parade, Parkville Department of Medicinal Chemistry Monash University Melbourne, Victoria 3052 Australia Attn: Matthew L Dennis
Billing Type	Invoice Matthew L Dennis 381 Royal Parade, Parkville Department of Medicinal Chemistry Monash University Melbourne, Australia 3052 Attn: Matthew L Dennis
Billing Address	
Total	0.00 AUD

This Agreement between Matthew L Dennis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	4006310786827
License date	Dec 12, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Drug Discovery
Licensed Content Title	A decade of fragment-based drug design: strategic advances and lessons learned
Licensed Content Author	Philip J. Hajduk and Jonathan Greer
Licensed Content Date	Mar 1, 2007
Licensed Content Volume	6
Licensed Content Issue	3
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 5
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Inhibitors of the folate pathway enzymes HPPK and DHPS
Expected completion date	Dec 2016
Estimated size (number of pages)	191
Requestor Location	Matthew L Dennis 381 Royal Parade, Parkville Department of Medicinal Chemistry Monash University Melbourne, Victoria 3052 Australia Attn: Matthew L Dennis
Billing Type	Invoice Matthew L Dennis 381 Royal Parade, Parkville Department of Medicinal Chemistry Monash University Melbourne, Australia 3052 Attn: Matthew L Dennis
Billing Address	
Total	0.00 AUD

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.
2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version. Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.
3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).
4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.
5. The credit line should read:
 Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
 For AOP papers, the credit line should read:
 Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

Note: For republication from the *British Journal of Cancer*, the following credit lines apply.

Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication) For AOP papers, the credit line should read:
 Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer*, the following credit line applies.

Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit <http://www.macmillanmedicalcommunications.com> for more information. Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

Note: For translation from the *British Journal of Cancer*, the following credit line applies.

Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.

Special Terms:

v1.1