Biotin protein ligase as a novel antifungal drug target

Louise Sternicki
B.Sc (Hons)
The University of Adelaide

A thesis to be submitted to the University of Adelaide, Australia for the degree of Doctor of Philosophy

THE UNIVERSITY of ADELAIDE

January 2020

School of Biological Sciences
Discipline of Biochemistry
The University of Adelaide
South Australia, Australia

Table of Contents

LIST OF INCLUDED PUBLICATIONS

The author acknowledges the copyright of published works contained within this thesis including:

Chapter 1:

Sternicki, L. M., Wegener, K. L., Bruning, J. B., Booker, G. W. & Polyak S. W. (2017) Mechanisms Governing Precise Protein Biotinylation. *Trends in Biochemical Sciences,* **42**(5): 383-394.

Chapter 3:

Satiaputra, J., Sternicki, L. M., Hayes, A. J., Pukala, T. L., Booker, G. W., Shearwin, K., Polyak, S. W. (2019) Native mass spectrometry identifies an alternate DNA-binding pathway for BirA from *Staphylococcus aureus*. *Scientific Reports,* **9**

Chapter 4:

Sternicki, L. M., Nguyen, S., Pendini, N. R., Wilce, M. C., Barran, P., Pacholarz, K., Pukala, T. L., Booker, G. W., Wegener, K. L., Polyak, S. W. (2019) Characterisation of class III biotin protein ligases; novel antifungal targets for human and agricultural pathogens. *Manuscript in preparation.*

Chapter 5:

Sternicki, L. M., Pukala, T. L., Pacholarz, K., Barran, P., Booker, G. W., Wegener, K. L., Polyak, S. W. (2019) Structural investigations of the N-terminal domain of the class III *Saccharomyces cerevisiase* Biotin Protein Ligase. *Manuscript in preparation.*

Appendix 1:

Thompson, A. T., Sternicki, L. M., Wegener, K. L., Lu, W., Zhu, L., Booker, G. W., Polyak, S. W. & Li, Y. (2016) Biotin Biology as a Target for New Anti-Tuberculosis Drugs. *Jiangsu Journal of Preventative Medicine,* **27**(3): 257-261.

Appendix 2:

Hayes, A. J., Satiaputra, J., Sternicki, L. M., Paparella, A. S., Feng, Z., Rodriguez, B. B., Feng, J., Teiu, W., Abell, A., Eijelkamp, B., Shearwin, K. E., Pukala, T. L., Booker, G. W., Polyak, S. W. (2019) Probing the mechanism of action and resistance of BPL inhibitors in

Staphylococcus aureus using an antibacterial sulfonyl-based mimic of the reaction intermediate, biotinyl-5′-AMP. *Manuscript in preparation.*

Authorisation to publish each paper has been given and provided in print for each chapter containing copyright and co-authored work, including acknowledgement of contribution to the work from each author.

Abstract

Pathogenic fungi are responsible for causing significant disease in both humans and plants. Fungal infections are responsible for 1.5 million deaths annually, and despite the availability of antifungal treatments, mortality rates remain high. Fungi not only cause human disease, but can also infect plants, crops and fruits, thereby, being a major threat for secure food production and the economy. Fungi are developing resistance to the few classes of antifungals that are available to treat human or agricultural infections. Hence, novel antifungal agents are required.

Biotin is an essential co-factor required for all forms of life. It is covalently attached to biotin-dependent enzymes and participates in the metabolic reactions these enzymes perform. The enzyme responsible for the ATP-dependent attachment of biotin onto these enzymes is biotin protein ligase (BPL), which attaches biotin onto a conserved lysine residue in the biotin domain of the biotin-dependent enzymes.

BPL is essential for the viability of bacteria and fungi, therefore, it is proposed as a promising target for the development of novel anti-infectives. Whilst inhibitors against the BPLs from *Staphylococcus aureus* and *Mycobacterium tuberculosis* have been reported, BPL has not been targeted for the development of novel antifungals. This thesis investigates the BPLs from the model fungi *Saccharomyces cerevisiae* and from two pathogenic fungi responsible for causing agricultural disease. Fungal BPLs are structurally unique compared to bacterial homologs as they contain an additional N-terminal extension prior to the common catalytic domain and C-terminal cap found in all BPLs. BPLs with this extension are termed class III BPLs. Currently there is no atomic resolution structure of a class III enzyme and the function of the additional N-terminal extension is unknown. This thesis investigates the structure of these class III enzymes.

The BPLs from *Saccharomyces cerevisiae*, *Botrytis cinerea* and *Zymoseptoria tritici* were characterised for their potential as antifungal drug targets. These three enzymes had similar overall structures as expected from their sequence homology. However, they displayed different stabilities and variable Michaelis constants for the substrates biotin, MgATP and the biotin domain, which implied subtle structural differences between the enzymes. These class III BPLs varied in their preferences for different species' biotin domains, providing further evidence for a 'substrate selectivity verification' mechanism whereby the N-terminal extension

of class III BPLs is proposed to distinguish between substrates and select only those appropriate for biotinylation. Finally, potent and selective *in vitro* inhibition of these class III BPLs was demonstrated using chemical analogs of the reaction intermediate, supporting the hypothesis that fungal BPLs are druggable antifungal targets.

The structure of the *S. cerevisiae* BPL (*Sc*BPL) was investigated in greater detail, and structural changes that occur upon ligand binding were examined. Whilst gross structural changes did not occur concomitant with ligand binding, smaller structural changes including rigidification and a reduction in the dynamic movement of the enzyme occurred such that *Sc*BPL stability increased. Homology modelling and hydrogen-deuterium exchange (HDX) mass spectrometry revealed a structured domain present in the N-terminal extension, with a predicted fold homologous to a glutamine amidotransferase. This N-terminal domain was linked to the structurally conserved BPL catalytic domain by a 160-residue linker that was mostly folded and/or buried. Together, HDX and homology modelling revealed that ligand binding induced local structural rearrangements in the N-terminal domain and the ligand-binding site of the catalytic domain. Finally, HDX also identified potential interaction surfaces of each domain indicating intramolecular interactions that govern BPL function.

Overall, this project aims to understand the structure of a class III BPL in order to target these enzymes for the development of novel antifungal agents for the treatment of both agricultural and clinical infections.

Declaration for thesis containing published work and/or work prepared for publication

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Louise Sternicki

26th July 2019

Acknowledgements

I would like to thank all of the people who have helped me over the past 4 years both with my PhD and beginning a career in medical science research.

Firstly, thank you to Grant Booker for taking me on as an Honours and then PhD student. Your insights and discussions have been very helpful not only for my project but my career in general. Thank you to Steven Polyak for supervising me during both my Honours and PhD. Thank you for your guidance, practical lab training and for teaching me about many aspects of a career in science. A big thank you to Kate Wegener for being my supervisor, mentor and friend. I have appreciated your support, scientific ideas, and career and life advice, often given over a chicken laksa lunch. I would also like to thank the other academics around Adelaide University that have assisted me throughout my PhD, both scientifically and with mentorship and advice. This especially includes Tara Pukala for being a surrogate supervisor and introducing me to the world of mass spectrometry. I have really appreciated your assistance, guidance, knowledge, support and mentorship. A thank you is also due to John Bruning for being a collaborative group leader and offering further insights and advice.

I would like to acknowledge many of the technical staff and facility managers for their assistance with instrument usage, maintenance and experiment optimisation. Without all of you things would have been a lot harder. This includes Phil Clements (NMR and MS facility), the Flinders MS facility, the Adelaide Proteomics Centre and the MLS stores and technical support unit.

A big thank you to Perdita Barran and the University of Manchester for hosting me for a research placement for two months in 2018. It was a fantastic opportunity to experience a different laboratory and collect novel data for my PhD. I would also like to thank Kamila Pacholarz for assisting with the experiments, and everyone in the lab for making me feel welcome and helping me during this time.

I would like to acknowledge all the member of the Booker and Bruning groups over the years I have been involved in the group. Thank you for the friendships and for bringing a social life to an otherwise very stressful and challenging environment. I would also like to thank all the friends and colleagues in the various labs across the biological sciences and chemistry departments.

I would like to give a special acknowledgement to the ASBMB for granting me an ASBMB Fellowship and the Fred Collins Award. This provided me with the financial means to undertake an amazing opportunity to complete research towards my PhD at the University of Manchester in the UK.

I also need to thank all of my friends and family that have supported and encouraged me during my PhD. Thank you to Emily for your support and for being there to listen and give advice when I have needed it. I have enjoyed the trips to Port Augusta or Mildura to visit and get away from the study for a short while. I would also like to thank my Mum and Dad for all of

your help, support and advice over the last few years. Thank you for your encouraging words to keep going when it was all too hard, and for your advice and consoling when my doubts were too great to ignore. I cannot thank you enough for everything you have done to help me through my PhD and to get to where I am now.

Finally, thank you to Richard for all your support and sacrifices during this process. Thank you for sticking with me through the good, the bad and the ugly of a PhD, and for being a constant source of support and encouragement.

Communications and Presentations

L. M. Sternicki, K. Pacholarz, P. Barran, T. L. Pukala, G. W. Booker, K. L. Wegener, S. W. Polyak (2019) Mass spectrometry approaches for the structural characterisation of class III Biotin Protein Ligases: novel antifungal drug targets. *44th Lorne Conference on Protein Structure and Function, Lorne, Victoria, Australia.* Poster Presentation. **Awarded** Student Poster Prize ($100).

L. M. Sternicki, S. Nguyen, T. L. Pukala, N. R. Pendini, S. Beckham, M. C. Wilce, G. W. Booker, K. L. Wegener, S. W. Polyak (2018) Structural and kinetic characterisation of class III Biotin Protein Ligases; novel antifungal drug targets. *ComBio Conference, Sydney, New South Wales, Australia.* Oral Presentation. **Awarded** ASBMB Fred Collins Fellowship ($3500).

L. M. Sternicki, T. L. Pukala, G. W. Booker, K. L. Wegener, S. W. Polyak (2018) Investigating the structure of class III Biotin Protein Ligases to understand biotinylation specificity. *43rd Lorne Conference on Protein Structure and Function, Lorne, Victoria, Australia.* Poster Presentation.

L. M. Sternicki, T. L. Pukala, G. W. Booker, K. L. Wegener, S. W. Polyak (2017) Structural studies to investigate the 'substrate verification' activity of class III Biotin Protein Ligases. *ComBio Conference, Adelaide, South Australia, Australia.* Oral Presentation. **Awarded** ASBMB Student Bursary ($250).

L. M. Sternicki (2017) New drugs to fight fungi. *The University of Adelaide School of Biological Sciences 10th Annual Postgraduate Research Symposium, Adelaide, South Australia, Australia.* Oral Presentation, Three-Minute Thesis Heats.

L. M. Sternicki, T. L. Pukala, G. W. Booker, K. L. Wegener, S. W. Polyak (2017) Structural studies to investigate the 'substrate verification' activity of class III Biotin Protein Ligases. *Adelaide Protein Group Student Awards Meeting, Adelaide, South Australia, Australia.* Poster Presentation.

L. M. Sternicki, T. L. Pukala, N. R. Pendini, M. C. Wilce, G. W. Booker, K. L. Wegener, S. W. Polyak (2016) Investigating the role of the N-terminal domain unique to class III Biotin Protein Ligases. *Australian Physiological Society & Australian Society of Biophysics Meeting, Adelaide, South Australia, Australia.* Oral Presentation.

L. M. Sternicki, T. L. Pukala, N. R. Pendini, M. C. Wilce, G. W. Booker, K. L. Wegener, S. W. Polyak (2016) Investigating the role of the N-terminal domain unique to class III Biotin Protein Ligases. *EMBL Australia Postgraduate Symposium, Adelaide, South Australia, Australia.* Poster Presentation.

L. M. Sternicki, T. L. Pukala, N. R. Pendini, M. C. Wilce, G. W. Booker, K. L. Wegener, S. W. Polyak (2016) Proof-reading ensures precise biotinylation by class III Biotin Protein Ligases.

The University of Adelaide School of Biological Sciences 9th Annual Postgraduate Research Symposium, Adelaide, South Australia, Australia. Poster Presentation. **Awarded** 1st Place Poster Prize.

N. R. Pendini, L. M. Sternicki, D. Cini, K. M. Downey, T. L. Pukala, S. Beckham, G. W. Booker, M. C. Wilce, S. W. Polyak (2016) New insights into the specificity of protein biotinylation. *Australian Society for Medical Research SA Division Scientific Meeting, Adelaide, South Australia, Australia.* Poster Presentation.

N. R. Pendini, L. M. Sternicki, D. Cini, K. M. Downey, T. L. Pukala, S. Beckham, G. W. Booker, M. C. Wilce, S. W. Polyak (2016) New insights into the specificity of protein biotinylation. *41st Lorne Conference on Protein Structure and Function, Lorne, Victoria, Australia.* Poster Presentation.

N. R. Pendini, L. M. Sternicki, D. Cini, K. M. Downey, T. L. Pukala, S. Beckham, G. W. Booker, M. C. Wilce, S. W. Polyak (2015) New insights into the specificity of protein biotinylation. *ComBio Conference, Melbourne, Victoria, Australia.* Poster Presentation.

Co-authored Communications and Presentations

Presenting authors underlined

A. J. Hayes, J. Satiaputra, A. S. Paparella, Z. Feng, L. M. Sternicki, B. Blanco-Rodriguez, A. D. Abell, G. W. Booker, S. W. Polyak (2016) Characterisation of resistance mechanisms to biotin protein ligase inhibitors, a novel antibacterial class targeting *Staphylococcus aureus. EMBL Australia Postgraduate Symposium, Adelaide, South Australia, Australia.* Oral Presentation.

A. S. Paparella, J Feng, B. Blanco-Rodriguez, L. M. Sternicki, M. A. Cooper, G. W. Booker, A. D. Abell, S. W. Polyak (2016) Template guided synthesis provides a route to new antibacterials and antifungals. *Gordon Research Seminar on Enabling New Antibiotic Discovery and Development, Tuscany, Italy.* Oral Presentation.

A. S. Paparella, J. Feng, B. Blanco-Rodriguez, L. M. Sternicki, M. A. Cooper, G. W. Booker, A. D. Abell, S. W. Polyak (2016) Template guided synthesis provides a route to new antibacterials and antifungals. *41st Lorne Conference on Protein Structure and Function, Lorne, Victoria, Australia.* Poster Presentation.

A. S. Paparella, J. Feng, B. Blanco-Rodriguez, M. Y. Yap, A. Hayes, D. J. Heim, L. M. Sternicki, M. A. Cooper, M. C. Wilce, G. W. Booker, A. D. Abell, S. W. Polyak (2015) Structure guided design of small molecule inhibitors of biotin protein ligase for antibiotic drug discovery. *Adelaide Pharmacology Group Student Meeting, Adelaide, South Australia, Australia.* Oral Presentation.

A. S. Paparella, J. Feng, B. Blanco-Rodriguez, L. M. Sternicki, M. A. Cooper, G. W. Booker, A. D. Abell, S. W. Polyak (2015) Template guided synthesis provides a new route to antibacterials and antifungals. *ComBio Conference, Melbourne, Victoria, Australia.* Oral Presentation.

Abbreviations

1D	one dimension
2D	two dimensions
3D	three dimensions
Å	Angstroms
$Å^2$	Angstroms squared
ACC	acetyl Co-A carboxylase
ACC1	acetyl Co-A carboxylase isoform 1
ACC2 (HFA)	acetyl Co-A carboxylase isoform 2
ADP	adenosine diphosphate
aIM-MS	activated ion mobility-mass spectrometry
AmAc	ammonium acetate
AMP	Adenosine monophosphate
Apo	unligand enzyme
ATP	Adenosine triphosphate
BCCD	biotin carboxyl carrier domain
BCCP	biotin carboxyl carrier protein
*Bc*BPL	*Botrytis cinerea* biotin protein ligase
BIO2	biotin synthase
BIO3	7,8-diamino-pelargonic acid (DAPA) aminotransferase
BIO4	dethiobiotin synthetase
BIO5	7-keto-8-aminopelargonic acid (KAPA) transporter
BirA	biotin retention protein A (Biotin inducible repressor)
BME	β-mercaptoethanol or 2-mercaptoethanol
bp	base pair
BPL	biotin protein ligase
bpl or *bpl1*	biotin protein ligase gene
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
*Ca*BPL	*Candida albicans* biotin protein ligase
*Ca*PC115	*Candida albicans* pyruvate carboxylase biotin domain, from the most C-terminal 115 residues of pyruvate carboxylase isoform 1
CCS	collision cross section
CD	circular dichroism

CIU	collision induced unfolding
CIU-MS	collision induced unfolding-mass spectrometry
CO_2	carbon dioxide
Cryo-EM	cryogenic-electron microscopy
D_2O	deuterium oxide
Da	Dalton
DAPA	7,8-diaminopelargonic acid
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTB	dethiobiotin
DTT	dithiothreitol
*Ec*BPL	*Escherichia coli* biotin protein ligase
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
ESI	electrospray ionisation
ESI-MS	electrospray ionisation-mass spectrometry
FPLC	fast phase protein liquid chromatography
GATase	glutathione amidotransferase
GST	glutathione S-transferase
HCl	hydrochloric acid
H/D	hydrogen/deuterium
HDMS	high definition mass spectrometer
HDX	hydrogen deuterium exchange
HDX LC-MS	hydrogen deuterium exchange liquid chromatography-mass spectrometry
HDX-MS	hydrogen deuterium exchange-mass spectrometry
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFA (ACC2)	acetyl CoA carboxylase isoform 2, mitochondrial located
HLCS	holocarboxylase synthetase, biotin protein ligase in humans
Holo	ligand-bound enzyme
hr	hour/s
HSQC NMR	heteronuclear single quantum coherence nuclear magnetic resonance
IC_{50}	half maximal inhibitory concentration
IM	ion mobility
IMAC	immobilised metal ion affinity chromatography

IM-MS	ion mobility-mass spectrometry
IPTG	isopropyl-β-D-thiogalactoside
K	kelvin
KAPA	7-keto-8-aminopelargonic acid
K_D	dissociation constant
kDa	kilo-Dalton
K_i	inhibition constant
K_M	Michaelis-Menten constant
kV	kilovolt
LB	luria broth (also known as Lysogeny broth, Lennox broth or Luria-Bertani broth)
LC-MS	liquid chromatography-mass spectrometry
LiOAc	lithium acetate
M	molar
mA	milliampere
mbar	millibar
MCD	multiple carboxylase deficiency
MES	2-(N-morpholino)ethanesulfonic acid
MIC	minimum inhibitory concentration
Micro-ED	micro-electron diffraction
min	minute/s
Mg	magnesium
MgATP	magnesium adenosine triphosphate
$MgCl_2$	magnesium chloride
MHz	mega-hertz
MOPS	3-(N-morpholino)propanesulfonic acid
MRE	mean residual ellipticity (measured in deg.cm-2.dmol-1)
MS	mass spectrometry
m/sec	metres per second
*Mt*BPL	*Mycobacterium tuberculosis* biotin protein ligase
MW	molecular weight
MWCO	molecular weight cut off
m/z	mass to charge ration (i.e. mass divided by charge)
n-	nano
NaCl	sodium chloride

NaH_2PO_4	monosodium phosphate
NaOH	sodium hydroxide
nESI	nano-electrospray ionisation
nESI-MS	nano-electrospray ionisation-mass spectrometry
NiNTA	nickel-nitrilotriacetic acid
nm2	nanometres squared
NMR	nuclear magnetic resonance
OD_{xnm}	optical density at x nm wavelength
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PC	pyruvate carboxylase
PC1	pyruvate carboxylase isoform 1
PC2	pyruvate carboxylase isoform 2
PDB	protein data bank
PMSF	phenylmethylsulfonylfluoride
PLP	pyridoxal 5′-phosphate
PPi	pyrophosphate
RMSD	root mean square deviation
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
s (or sec)	second/s
*Sa*BirA	*Staphylococcus aureus* biotin retention protein A (same as *S. aureus* biotin protein ligase)
*Sa*BioO	*Staphylococcus aureus* biotin biosynthesis operon
*Sa*BioY	*Staphylococcus aureus* biotin transporter, BioY
*Sa*BPL	*Staphylococcus aureus* biotin protein ligase (same as *S. aureus* biotin retention protein A)
*Sa*PC90	*Staphylococcus aureus* pyruvate carboxylase biotin domain, from the most C-terminal 90 residues of pyruvate carboxylase
SAXS	small angle X-ray scattering
*Sc*BPL	*Saccharomyces cerevisiae* biotin protein ligase
*Sc*PC104	*Saccharomyces cerevisiae* pyruvate carboxylase biotin domain, from the most C-terminal 104 residues of pyruvate carboxylase isoform 1
SDS	sodium dodecyl sulphate

SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
spp.	species
SPR	surface plasmon resonance
TBS	tris buffered saline
TCA	tricarboxylic acid
T_M	melting temperature
ToF	time of flight
Tris	Tris(hydroxymethyl)aminomethane
TWIM-MS	travelling wave ion mobility-mass spectrometry
UA	urea amidolyase
UV	ultraviolet
V_{50}	inflection point of Boltzmann sigmoidal curve
V	volts
VHT1	vitamin H transporter 1, biotin transporter for *Saccharomyces cerevisiase*
WT	wild-type
x g	x g-force (also known as relative centrifugal force)
YP	yeast extract peptone dextrose (also known as YEPD or YPD)
*Zt*BPL	*Zymoseptoria tritici* biotin protein ligase

CHAPTER 1:

INTRODUCTION LITERATURE REVIEW

CHAPTER OUTLINE

This chapter is a published review outlining prior literature about the enzyme Biotin Protein Ligase (BPL), the main focus of this thesis. This review provides a background to BPLs, particularly focusing on eukaryotic BPLs, including those from fungi that contain a unique structure compared to other BPL structural classes. This chapter highlights current knowledge regarding eukaryotic BPLs and reveals information that is lacking and requires investigation in order to fully understand these enzymes.

STATEMENT OF AUTHORSHIP

Statement of Authorship

Title of Paper	Mechanisms Governing Precise Protein Biotinylation
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Published in Trends in Biochemical Science, May 2017, Vol 42, No. 5, 383-394. doi: 10.1016/j.tibs.2017.02.001

Principal Author

Name of Principal Author (Candidate)	Louise M Sternicki
Contribution to the Paper	Planned the outline of the manuscript, conducted the literature research, prepared the manuscript and figures
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 6/3/2019.

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Kate L Wegener
Contribution to the Paper	Preparation of the manuscript
Signature	Date 6/3/19

Name of Co-Author	John B Bruning
Contribution to the Paper	Preparation of the manuscript
Signature	Date 6/3/19

Name of Co-Author	Grant W Booker		
Contribution to the Paper	Preparation of the manuscript		
Signature		Date	6/3/2019

Name of Co-Author	Steven W Polyak		
Contribution to the Paper	Helped plan the outline and prepare the manuscript		
Signature		Date	6/3/2019

Trends in Biochemical Sciences

CellPress

Review

Mechanisms Governing Precise Protein Biotinylation

Louise M. Sternicki,[1] Kate L. Wegener,[1] John B. Bruning,[1] Grant W. Booker,[1] and Steven W. Polyak[1,*]

Protein biotinylation is a key post-translational modification found throughout the living world. The covalent attachment of a biotin cofactor onto specific metabolic enzymes is essential for their activity. This modification is distinctive, in that it is carried out by a single enzyme: biotin protein ligase (BPL), an enzyme that is able to biotinylate multiple target substrates without aberrant-off target biotinylation. BPL achieves this target selectivity by recognizing a sequence motif in the context of a highly conserved tertiary structure. One structural class of BPLs has developed an additional 'substrate verification' mechanism to further enable appropriate protein selection. This is crucial for the precise and selective biotinylation required for efficient biotin management, especially in organisms that are auxotrophic for biotin.

Protein Biotinylation: An Important Post-Translational Modification Requiring Tight Regulation

Post-translational modification of proteins occurs in all living organisms. It governs many important cellular processes including signaling, protein localization and trafficking, the activation of enzymes, and protein degradation. The enzymes responsible for post-translational modifications often recognize a short, linear amino acid sequence, or 'recognition motif', present within the protein substrate targeted for chemical alteration. For example, kinases recognize specific sequences immediately surrounding the hydroxyl-containing residue that is ultimately phosphorylated [1,2]. Likewise, methyltransferases, responsible for histone modification, recognize lysine residues in the flexible tails of histones [3,4]. By contrast, the attachment of **biotin** (see Glossary) onto protein relies on the target lysine being housed within a structured domain typically composed of 60–80 amino acids. This is unusual because post-translational modification recognition motifs generally do not require a structural context, with one other notable exception being lipoylation [5]. Biotin is widely utilized in biotechnological applications as a labeling reagent or affinity purification tag owing to its extremely high affinity for (strept)-avidin. However, the underlying biology of this natural modification in organisms is seldom appreciated.

All living organisms have between one and six **biotin-dependent enzymes**, highlighting biotin as an essential micronutrient in all kingdoms of life (reviewed in [6]). Protein biotinylation involves the precise attachment of biotin (also known as vitamin B7 or H) onto a single lysine residue present on the **biotin carboxyl carrier domain** (BCCD) of specific metabolic enzymes. These biotin-dependent enzymes play key roles in important metabolic pathways such as lipogenesis, gluconeogenesis, amino acid metabolism, and energy transduction, among others [6]. When attached to a BCCD, the biotin cofactor facilitates the transfer of a carboxyl group between metabolites during carboxylation, decarboxylation, and transcarboxylation reactions. Without biotin these enzymes are unable to perform their vital metabolic functions. The sole enzyme responsible for modifying all the biotin-dependent enzymes in the

Trends

BPL catalyzes the exquisitely precise attachment of biotin onto protein substrate(s).

Structural biology has helped to provide powerful new insights into the molecular basis of protein biotinylation.

BPL recognizes a consensus AMKM motif in the protein substrate within the context of a flattened β-barrel tertiary structure.

This tertiary structure is conserved amongst all biotin-dependent enzymes. This allows one enzyme, BPL, to modify multiple substrates within one organism, and without erroneously biotinylating off-target proteins.

Mammalian BPLs can discriminate between protein substrates. This allows these biotin auxotrophs to effectively utilize the precious micronutrient.

A unique N-terminal extension on mammalian BPLs enhances catalytic activity by providing stabilizing interactions with substrates. It is proposed that the interactions generated by this domain may allow 'substrate verification' to select appropriate proteins for modification.

[1]School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, Australia

*Correspondence: steven.polyak@adelaide.edu.au (S.W. Polyak).

Trends in Biochemical Sciences

CellPress

cell is **biotin protein ligase** (BPL) [7]. BPL shows exceptional specificity in selecting appropriate substrates to modify, with no aberrant off-target biotinylation. Biochemical and structural studies are aiding our understanding of how this precise post-translational modification is achieved.

A variety of strategies are employed by organisms to fulfill their biotin requirements. Some bacteria, fungi, and plants synthesize biotin *de novo* [8]. Other organisms use high-affinity vitamin transporters to scavenge exogenous biotin from their environment [9]. Bacteria such as *Escherichia coli*, *Bacillus subtilis,* and *Staphylococcus aureus* possess both importer and biosynthetic enzymes, and can regulate the expression of these enzymes depending on biotin supply and cellular demand [10,11]. By contrast, mammals obtain this micronutrient solely from their environment, primarily from intestinal microflora and dietary sources, and from recycling of degraded biotin-dependent enzymes [12–14]. Accordingly, biotin auxotrophs require tight control over biotin utilization to ensure thrifty use of this precious micronutrient. Therefore, it is proposed that mammalian BPLs have developed specific mechanisms to recognize and verify appropriate substrates to avoid inappropriate biotinylation. The structure of BPLs, and the mechanisms that govern specificity, form the basis of this review.

Structural Classes of BPL

All BPLs can be classified into three distinct structural classes (Figure 1A). The class I BPLs contain two domains – an SH2-like catalytic domain containing the biotin and ATP binding sites, and an SH3-like C-terminal cap [15–21]. This SH2:SH3 fusion containing the minimal requirements for protein biotinylation is highly conserved in all BPLs. Class II BPLs contain this conserved module with an additional N-terminal winged helix-turn-helix motif to enable DNA binding [22–25]. These bifunctional proteins serve as both biotin ligases and biotin-dependent transcription factors. Transcriptional regulation of biotin synthesis and transport by the protein required for biotin utilization allows these bacteria to finely balance cellular demand with intracellular biotin concentration and environmental supply [11,26]. BPLs belonging to classes I and II have been well characterized, and crystal structures of several examples have now been reported, including the class I proteins from *Mycobacterium tuberculosis* [15–18], *Pyrococcus horikoshii* [19,20], and *Aquifex aeolicus* [21], and of the class II BPLs from *E. coli* [22–24] and *S. aureus* [25]. Therefore, most of the current structural knowledge about protein biotinylation has been derived from bacterial and archaeal enzymes. Plant BPLs can be considered as class I enzymes with additional N-terminal organelle-targeting sequences that allow isoforms to be differentially localized in plant cells [27,28]. Class III BPLs, that are present in eukaryotes such as yeast, insects, and mammals, also contain an N-terminal extension in addition to the conserved catalytic module [29]. This N-terminal region is much larger than the class II extension, contains no predicted DNA-binding motifs, and shares no homology with the N-terminal domain of class II family members. Currently no high-resolution structural data have been reported for a class III BPL catalytic domain, nor for the N-terminal extension. As will be described later, it has become apparent that the N-terminal region provides a 'substrate verification' activity to ensure that only the correct proteins are selected for biotinylation.

Mechanism of Protein Biotinylation by BPL

Biotin is covalently attached onto a single target lysine residue in biotin-dependent enzymes. This key lysine is located within a highly conserved AMKM motif housed within the BCCD of the protein substrate. The structure of this protein substrate is described in detail later. BPL employs a two-step reaction requiring biotin and ATP to catalyze protein biotinylation (Figure 2A). The ligands bind in adjacent pockets in the enzyme such that the carboxylate group of biotin can ligate to the α-phosphate group of ATP to form the reaction intermediate, biotinyl-5′-AMP [30]. The BPL:biotinyl-5′-AMP 'holo' complex then forms a protein:protein

Glossary

Biotin: a member of the B class of water-soluble vitamins (B7). For historical reasons it is also known as vitamin H, from the German word *Haut*, meaning 'skin', because skin lesions caused by eating raw egg could be reversed by biotin-containing extracts from yeast and liver.

Biotin carboxyl carrier domain (BCCD): a highly conserved structure present in all biotin-dependent enzymes. This protein domain, sometimes referred to as the 'biotin domain', houses the sole lysine residue to which the biotin cofactor is covalently attached. The BCCD is connected to the biotin-dependent enzyme through a flexible linker that permits biotinyl-BCCD to oscillate between two partial reaction sites during carboxylation, decarboxylation, and transcarboxylation reactions. This 'swinging arm' mechanism allows CO_2 to be transferred between metabolites. For historical reasons, 'biotin carboxyl carrier protein' (BCCP) commonly appears in the literature because some of the earliest biotin-dependent enzymes studied were multisubunit examples where one protein subunit adopts the BCCD structure.

Biotin-dependent enzymes: a family of metabolic enzymes that catalyze carboxylation, decarboxylation, and transcarboxylation reactions using a covalently attached biotin prosthetic group. The biotin cofactor is essential for directly binding CO_2. Biotin-dependent enzymes are ubiquitous in all three kingdoms of life.

Biotin protein ligase (BPL): this enzyme (convention identifier 6.3.4.15) catalyzes the covalent attachment of biotin onto a single lysine residue present in the BCCD of the biotin-dependent enzymes. The BPL homolog in *Homo sapiens* is also known as holocarboxylase synthetase (HCS).

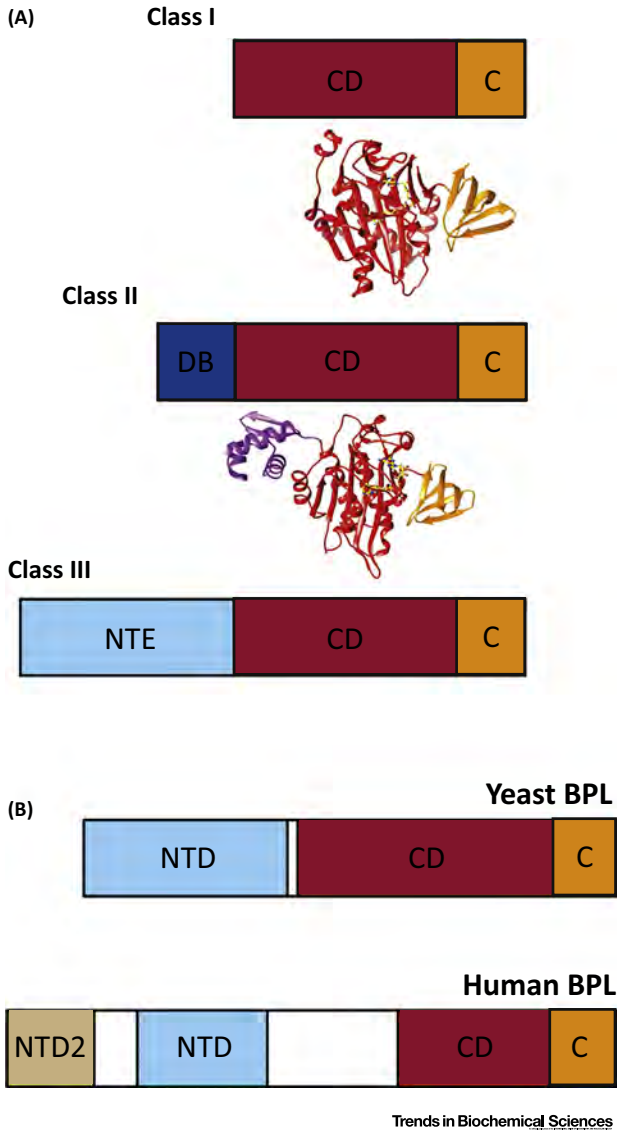


Figure 1. The Three Biotin Protein Ligase (BPL) Structural Classes with Examples of the Domain Architecture of Two Class III BPLs. (A) Class I BPLs contain the conserved catalytic domain (CD, dark red) and the C-terminal cap (C, orange), which constitute the minimal module necessary for biotinylation. Class II BPLs contain this essential module together with an N-terminal DNA-binding domain (DB, dark blue) that enables the BPL to transcriptionally regulate biotin supply. An N-terminal extension (NTE, light blue) is also present in class III BPLs that bears no homology to the class II DNA-binding domains. Examples of crystal structures from each of these classes are provided (*Mycobacterium tuberculosis* (class I) (PDB: 4OP0) [15], *Staphylococcus aureus* (class II) (PDB: 4DQ2) [34]), except for class III for which there are no atomic resolution structures. (B) A structured N-terminal domain (NTD, light blue) is located within the extension of the *Saccharomyces cerevisiae* and human class III BPLs [33,59,65]. This domain enhances catalytic activity by interacting with the biotin carboxyl carrier domain (BCCD) substrate [33,67]. The human BPL also contains another N-terminal domain (NTD2, tan) encompassing residues 1–100 that is important for substrate recognition [71].

interaction with the BCCD substrate, precisely positioning the target lysine such that its ε-amino group can pick up the biotin in a nucleophilic substitution reaction, releasing AMP [19,22,25].

Most characterized BPLs exhibit an ordered binding mechanism whereby biotin binding precedes that of ATP [31]. Biotin induces the ordering of a biotin-binding loop (residues 110–128 in *E. coli*, shown in Figure 2B) [22] to reposition a key tryptophan residue (W123) within the active site, allowing the nucleotide triphosphate to bind through a Π–Π stacking interaction with the adenine ring of ATP [24,25]. The biotin-binding loop closes over the active site and is stabilized through a salt-bridge interaction between R118 (in the biotin-binding loop) and D176. A highly conserved lysine at position 183 in *E. coli* BPL (K111 in *P. horikoshii*) is proposed to facilitate the synthesis of biotinyl-5′-AMP by aiding the chemical ligation between the carboxyl group of biotin and the α-phosphate of ATP, as well as balancing the negative charge of the reaction intermediate [20]. An additional hydrophobic loop (residues 212–223 in *E. coli*, the 'adenylate-binding loop') becomes ordered and

Trends in Biochemical Sciences

CellPress

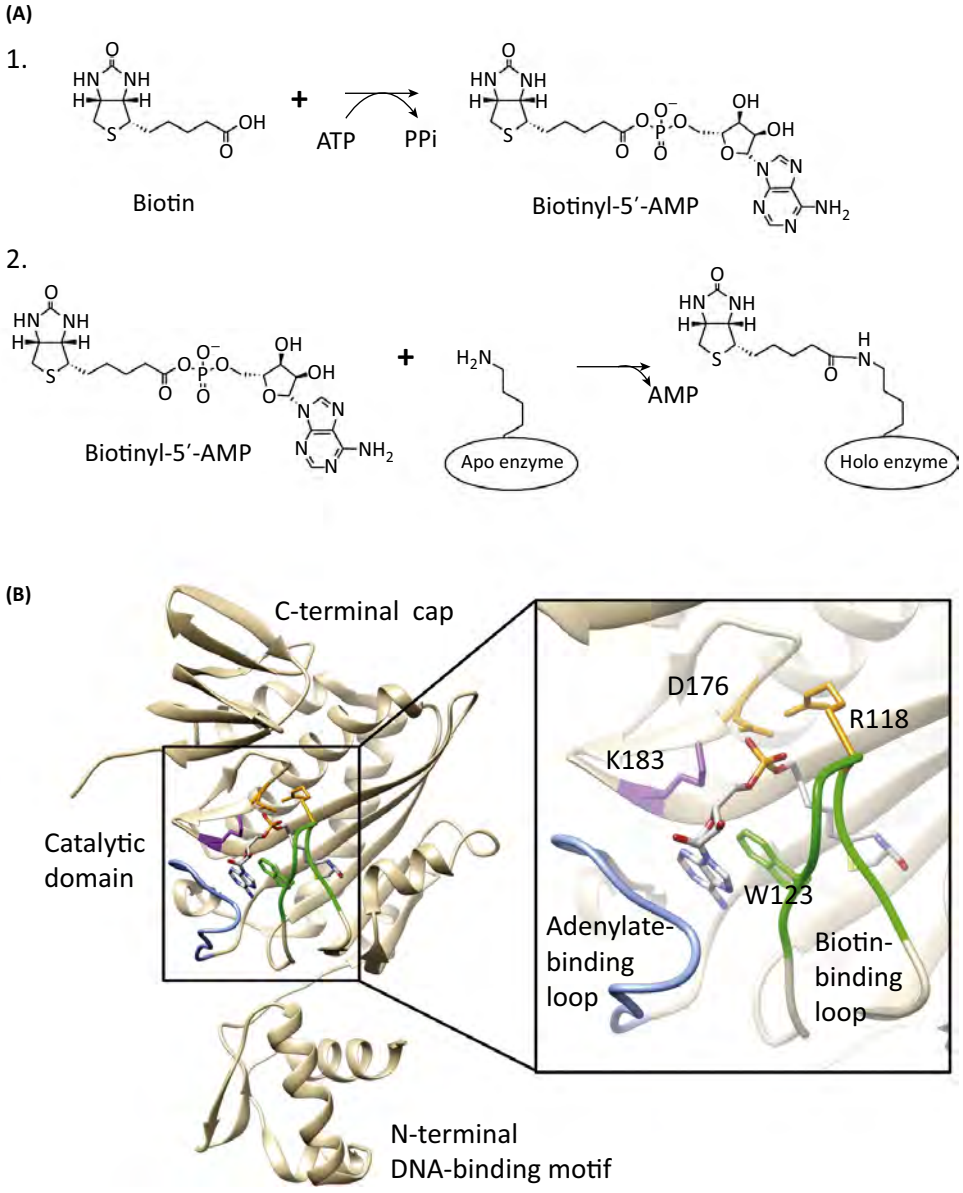


Figure 2. Synthesis and Binding of the Biotin Protein Ligase (BPL) Reaction Intermediate. (A) The two-step reaction mechanism employed by BPL for protein biotinylation. (1) Biotin and ATP condense to form the reaction intermediate, biotinyl-5′-AMP. (2) A key lysine housed in the biotin carboxyl carrier domain (BCCD) substrate is selected for the covalent attachment of biotin, with the release of AMP. (B) *E. coli* BPL (PDB: 2EWN) [24] in complex with biotinyl-5′-AMP (grey). Biotin first binds to BPL, causing restructuring of the previously disordered biotin-binding loop (green) that inhibits the dissociation of biotin. Interactions between the side chains of R118 and D176 (both orange) help to stabilize this closed conformation in which the biotin-binding loop folds over the active site. W123 (side chain shown) is also repositioned as a result of this conformational change, thereby allowing ATP to bind. ATP binding results in the ordering of the adenylate-binding loop (blue) that further stabilizes the binding of the reaction intermediate to prevent dissociation. This mechanism results in the ordered binding of biotin before ATP. A key lysine residue (K183, purple), located within the active site, is required for the catalytic ligation of biotin and ATP to form the reaction intermediate biotinyl-5′-AMP.

encases the adenyl group to further stabilize binding of the reaction intermediate [32]. This ordered binding mechanism is present in BPLs from all three classes [15,24,33–35]. Notable exceptions are the BPLs from the thermophilic organisms *P. horikoshii* and *A. aeolicus* that can bind either ligand independently. Crystal structures of these BPLs reveal that the biotin- and adenylate-binding loops are pre-ordered such that the enzyme is competent to bind either ligand first [20,21].

Both the biotin- and adenylate-binding loops play a key role in shielding the labile reaction intermediate from hydrolysis and preventing inappropriate reaction with non-target substrates. Mutation of R118 in *E. coli* BPL results in heightened dissociation of biotinyl-5′-AMP from the active site such that it can react with other proteins in the vicinity of the mutant enzyme [36,37]. Thus, controlled release of the reaction intermediate from BPL distinguishes precise from promiscuous biotinylation. The non-specific biotinylation activity of R118G mutant BPL has been exploited for proximity labeling of proteins that can be analyzed through streptavidin-based capture techniques and mass spectrometry [38].

Recognition of the BCCD As the Substrate for Biotinylation

Once the holo-BPL complex is formed, it is receptive to protein:protein interactions with an appropriate BCCD substrate. There is high structural conservation between the BCCDs of different biotin-dependent enzymes, as well as between species orthologs [29], as evidenced by the finding that many BPLs can biotinylate substrates from non-cognate species [39,40]. BCCDs contain all the requirements necessary for BPL recognition, association, and biotinylation [39,41,42]. Biotinylation is therefore a distinctive post-translational modification because the natural substrates do not possess a consensus sequence that BPL recognizes, but instead contain a sequence in a defined spatial arrangement. Despite this, a short unstructured 14-residue peptide is specifically biotinylated by only *E. coli* BPL [43,44]. However, this ‘avitag’ is not a natural substrate, and is instead a tool developed for various biotechnology applications, and only the recognition of natural substrates will be discussed here.

The structures of many BCCDs have been reported since the first examples from *E. coli* acetyl-CoA carboxylase were solved by both X-ray crystallography [45] and nuclear magnetic resonance (NMR) [46]. All BCCDs consist of the same basic structure: a flattened β-barrel composed of two four-stranded β-sheets (Figure 3A). The N- and C-termini are positioned together at one end of the structure, with the biotin-accepting lysine being located at the opposite end. This lysine is located within the conserved AMKM motif on a highly exposed tight hairpin loop between β-strands 4 and 5 [45]. Truncations of the *E. coli* domain demonstrate the smallest recognizable motif for biotinylation contains at least 35–40 residues on either side of the target lysine [45–49]. These flanking residues are responsible for maintaining the folded structure, stabilized by a hydrophobic core. Hence, BPL only recognizes the AMKM motif in the context of the folded BCCD.

Only one BPL:BCCD complex has been crystallized, namely the class I BPL from the thermophile *P. horikoshii* with its cognate BCCD (Figure 3B) [19]. This structure was solved in the presence of biotin and ATP to promote formation of the complex. Key BPL residues R48 and K111 (equivalent to R118 and K183 in *E. coli*) were mutated to disrupt binding of the triphosphate moiety of ATP to prevent reaction catalysis [19]. The structure reveals that the BPL and the BCCD interact to form an extended intermolecular β-sheet with strands contributed from both binding partners (Figure 3B) [19]. The BPL contacts are localized around the β2-strand of the BCCD and its preceding loop, together with β-strands 4 and 5, and the intervening hairpin that contains the target lysine. Binding of the substrate positions the target lysine side chain into the BPL active site, proximal to biotinyl-5′-AMP [19].

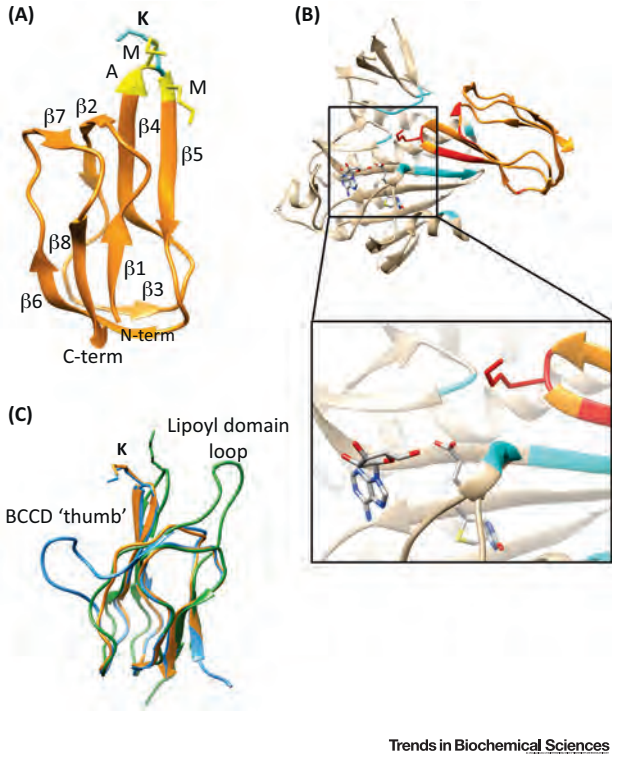


Figure 3. Structure of the Biotin Carboxyl Carrier Domain (BCCD) and Its Interaction with Biotin Protein Ligase (BPL). (A) The BCCD from *Pyrococcus horikoshii* (orange) (PDB: 2EJG) [19] adopts a flattened β-barrel tertiary fold conserved in all BPL substrates. Side chains of the residues in the AMKM recognition motif (yellow) are highlighted; as is the lysine targeted for biotinylation (blue). (B) Crystal structure of the *P. horikoshii* BPL (tan) with its cognate BCCD (orange) from the *P. horikoshii* methylmalonyl-CoA decarboxylase (PDB: 2EJG) [19]. Regions of the BPL involved in hydrogen bonding interactions of the BPL:BCCD complex are shown in blue, and the BCCD regions in dark red [19]. (C) Overlay of BCCDs from *Pyrococcus horikoshii* (orange) (PDB: 2EJG) [19], *Escherichia coli* (blue) (PDB: 1BDO) [45], and a lipoyl domain from the 2-oxoglutarate dehydrogenase complex in *Azotobacter vineland II*) (green) (PDB: 1GHJ) [75]. All three structures adopt the same flattened β-barrel structural fold. Notable features are the thumb structure between β-strands 2 and 3 in the *E. coli* protein (blue) and the protruding loop between β-strands 1 and 2 in the lipoyl domain. These loops are suggested to aid recognition by the correct ligase [53,55,56]. Abbreviations: C-term, C terminus; N-term, N terminus.

BCCDs also share structural homology to lipoyl-accepting domains, with similar overall structural folds despite limited sequence homology (Figure 3C) [50,51]. Lipoylated enzymes, such as pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase that play key roles in oxidative and single carbon metabolism, undergo an analogous post-translational modification [5]. In these cases lipoyl protein ligase appends lipoic acid to a lysine contained in a DKA motif that is also located on a β-hairpin turn between β-strands 4 and 5 [52], similar to the biotinylation motif [50,53]. Despite structural conservation of the biotinyl- and lipoyl-carrier domains, and also the structural conservation between the modifying enzymes themselves [54], modifications only occur to the appropriate substrate. Sequence alignments reveal that lipoyl domains contain an enlarged loop between the β1 and β2 strands, which is proposed to prevent engagement of BPL [53]. The *E. coli* BCCD contains a thumb-like protrusion between β-strands 2 and 3 that prevents aberrant lipoylation (Figure 3B) [55]. However, the *E. coli* BCCD appears to be unique amongst BPL substrates in having this feature. Recently, a domain from *Bacillus subtilis* was discovered that can be both lipoylated and biotinylated when expressed in *E. coli* [56]. This ability to be modified by both ligases arises from the lack of a thumb between β-strands 2 and 3, and an extended loop between β-strands 1 and 2 [56]. This highlights that these protrusions define structural features to ensure the correct recognition of the appropriate substrate by the specific ligase in some bacteria and archaea.

The mechanisms of substrate differentiation in organisms with thumbless BCCDs remain unclear. Of course, differences in primary structure also play a role, and BPL recognizes the key lysine as well as its flanking hydrophobic residues. By contrast, lipoyl domains feature negatively charged residues within this motif. The lipoyl protein ligase also shows affinity for

residues in β-strand 5, including another exposed acidic residue at the +4 position with respect to the target lysine [50,55,57]. Thus, structural features of the domain combine with specific side-chain identities to influence the specificity of the BPL:substrate interaction to select only those appropriate for biotinylation.

Substrate Verification Capabilities of Class III BPLs

Much of our understanding about protein biotinylation has been derived from studies on the class I and II enzymes from archaea and bacteria. The class III enzymes have not been as extensively investigated, but nonetheless play an important role in the biology of biotin. The remainder of the review will focus upon research that has extended our knowledge regarding class III BPLs. Biotin auxotrophs, such as mammals, insects, and some fungi, must satisfy their biotin requirement by sequestering this vitamin from exogenous sources. Owing to this limited supply, biotin must be well managed. In higher organisms, cells often contain multiple biotin-dependent enzymes that reside in various cellular compartments. To aid efficient utilization of biotin, it is an advantage to (i) ensure precise biotinylation, and (ii) select between the allowed substrates as required. In humans the profound mismanagement of biotin, or congenital defects in BPL, results in juvenile multiple carboxylase deficiency (MCD) (Box 1), which can result in death if untreated [58]. The detrimental consequences of inefficient biotin management demonstrate the importance of the thrifty use of available biotin and precise substrate recognition by BPL. Structural studies have begun to better define class III BPLs and their additional mechanisms for appropriate protein selection through a substrate verification function.

Gravel and Campeau first proposed a BPL substrate verification activity that allows class III BPLs to discriminate between closely related substrates. They found that human BPL (also termed holocarboxylase synthetase, HCS) can differentiate between a bacterial BCCD and the human equivalent in propionyl-CoA carboxylase [59]. Since then, evidence has been mounting to suggest that class III BPLs are more selective in the substrates they biotinylate than are the class I and II equivalents. Ingaramo and Beckett, for example, compared representatives from each of the BPL classes and found the human class III BPL to be the most selective, with a strong preference for biotinylating its cognate BCCD [40]. By contrast, the archaeal class 1 BPL

Box 1. Biotin-Dependent Enzymes in Mammals and Multiple Carboxylase Deficiency

Mammals possess five biotin-dependent enzymes that play key roles in important metabolic pathways; acetyl-CoA carboxylase-1 (synthesis of malonyl-CoA for fatty acid synthesis), acetyl-CoA carboxylase-2 (synthesis of malonyl-CoA to regulate β-oxidation of fatty acid synthesis), pyruvate carboxylase (replenishment of tricarboxylic acid cycle with oxaloacetate, and gluconeogenesis in the liver), propionyl-CoA carboxylase (synthesis of methylmalonyl-CoA during metabolism of the amino acids valine, isoleucine, methionine, and threonine, as well as of odd-numbered fatty acid chains) and 3-methylcrotonyl-CoA carboxylase (synthesis of 3-methylglutaconyl-CoA during the catabolism of leucine). All five biotin-dependent enzymes are biotinylated through the activity of a single BPL, also known as holo-carboxylase synthetase (HCS) [64].

Failure to adequately manage and utilize biotin supply can result in the human neonatal syndrome multiple carboxylase deficiency (MCD). MCD results from reduced activity of all five biotin-dependent enzymes due to (i) inadequate biotin supply due to malnutrition, (ii) compromised activity of human BPL (HCS deficiency, OMIM 253270) [58], (iii) reduced activity of the recycling enzyme biotinidase (OMIM 253260) [13], or (iv) inability to absorb exogenous biotin through a high-affinity vitamin transporter [73]. Symptoms of MCD include ketoacidosis, feeding difficulties, hypotonia, seizures, developmental delay, and dermal abnormalities such as rashes, dryness of the skin, and alopecia [14].

MCD is rare but serious because the most severe symptoms can lead to coma and death if untreated [13,68,70,74]. The only available treatment option for patients is 10–20 mg of supplementary dietary biotin, which is well tolerated and effective in the majority of cases. This treatment overcomes many forms of the biochemical and clinical markers of diseases where biotin transport, recycling, or utilization are impeded. Most characterized mutations in human BPL map to the catalytic domain of the enzyme, reducing its affinity for biotin [29]. However, not all patients respond to this treatment and alternative treatments are unsatisfactory.

Trends in Biochemical Sciences

CellPress

from *P. horikoshii* biotinylated all the substrates used in the study, while the *E. coli* BPL (class II) biotinylated only its own and the human BCCDs [40], with similar *in vitro* enzyme kinetics [60]. Similarly, a sixfold increase in K_m and a threefold decrease in reaction rate were reported when human BPL utilized a thumb-containing *E. coli* BCCD as the substrate compared to a thumbless human substrate [61].

The substrate verification ability of class III BPLs potentially allows selection between the different biotin-dependent enzymes in a single organism to use biotin most efficiently. Human BPL has demonstrated such selectivity between the five carboxylases it modifies *in vitro*. Kinetic experiments revealed that human BPL associates with biotin-dependent enzymes present in the mitochondria at faster rates than for cytosolic substrates [62]. Thus, differential localization of BPL into the various cellular compartments containing the biotin-dependent enzymes may be one potential mechanism to control biotin utilization in eukaryotes [63,64].

Delineating Structure and Functional Aspects of Class III BPLs

Although an atomic resolution structure of a class III BPL remains elusive, biochemical and biophysical studies of the human and yeast *S. cerevisiae* BPLs have provided valuable new insights. Limited proteolysis studies revealed the yeast BPL contains a 27 kDa N-terminal domain connected to a 50 kDa catalytic domain through a protease-sensitive linker (Figure 1B) [65]. Although this 50 kDa catalytic domain shares homology with the catalytic domains of other BPLs, the N-terminal domain demonstrates highest structural homology with various glutamine amidotransferases and hydrolases (identified using the Phyre fold recognition server) [66]. Protease susceptibility was reduced following incubation with saturating concentrations of the substrates biotin and MgATP, suggesting that yeast BPL compacts upon ligand binding (Figure 4) [65]. Removal of the N-terminal domain (residues 1–233) reduced yeast BPL's biotinylation activity by 3500-fold, and this level of activity was not sufficient for *E. coli* to be viable in a genetic complementation assay. Increasing biotin concentrations did not restore catalytic activity, demonstrating that a reduced K_m for biotin was not responsible for the change in activity [65]. Therefore, the structured N-terminal domain characteristic of class III BPLs appears to enhance catalytic activity.

Limited proteolysis studies predict the presence of a similar stable domain between residues 160–314 in the N-terminal extension of human BPL (Figure 1B) [33]. Complementation assays using fragments of the enzyme recombinantly expressed in a strain of *E. coli* harboring a temperature-sensitive BPL have provided a genetic means to identify enzymatically active variants [59]. These studies supported the presence of a structured domain between residues 166–290 that was necessary for activity [59]. Sequence alignments revealed that residues 160–314 of human BPL share 23% similarity with the N-terminal domain of yeast BPL, and Phyre structural homology searches returned the same matches for this region as with the yeast BPL. This N-terminal domain is important for catalytic function because a construct with the domain deleted was unable to rescue haploid spores from a *bpl*$^-$ strain of *S. cerevisiae* [67]. Yeast two-hybrid experiments also revealed that the N-terminal domain interacts with the BCCD [67]. It is through this interaction with the protein substrate that the N-terminal domain is suggested to enhance BPL catalytic activity.

Multiple Carboxylase Deficiency Contributes to Understanding the Role of the N-Terminal Domain.

Studies on MCD-causing point mutations provide clues into the mechanism by which the N-terminal extension enhances BPL activity. Patients with the missense mutations L216R and L237P (mapping to the N-terminal domain) respond poorly, or not at all, to doses of up to 200 mg of biotin daily [68–70]. Therefore, the effect of these amino acid substitutions cannot

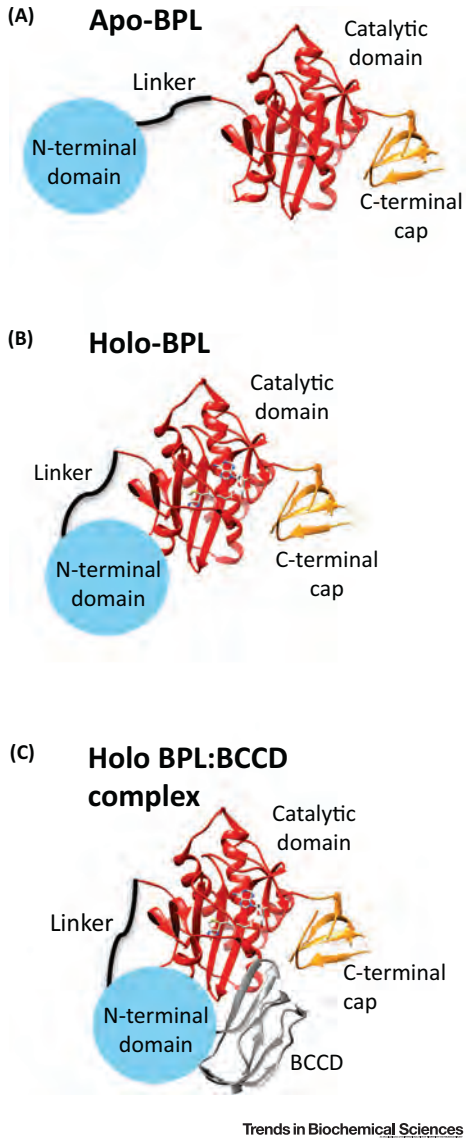


Figure 4. Proposed Model of the Interaction of the Class III Yeast Biotin Protein Ligase (BPL) with a Biotin Carboxyl Carrier Domain (BCCD) Substrate. (A) The N-terminal domain of the yeast BPL is separated from the catalytic domain by a protease-sensitive linker. (B) Limited proteolysis revealed that these domains come together upon ligand binding to constrain and prevent proteolysis of the linker region, suggesting that BPL undergoes a conformational change before interacting with the BCCD substrate [65]. (C) The N-terminal domain then interacts with the BCCD to stabilize this protein:protein interaction and allow biotinyl transfer to proceed. The structures used to model the BPL and BCCD are from the structure of the *P. horikoshii* complex (PBD: 2EJG) [19].

simply be due to decreased affinity for biotin. Surface plasmon resonance studies have shown that the L216R and L237P mutations reduce the affinity of human BPL for the BCCD of pyruvate carboxylase [33]. Association rates between immobilized BPL and the substrate were similar for the wild-type and the two missense mutant proteins. However, dissociation rates were 15-fold faster for the MCD mutants relative to the wild-type. This caused release of the BCCD before the biotinyl-transfer reaction could be completed [33]. Therefore, a key activity of the N-terminal domain of BPL is to stabilize the transient BPL:BCCD interaction long enough to permit the transfer of biotin from biotinyl-5′-AMP onto the protein substrate.

Trends in Biochemical Sciences

CellPress

Human BPL Contains an Additional N-Terminal Domain with Potential Catalytic and Recognition Functions

Computational structural analysis predicted an additional domain in the first 100 amino acids of human BPL that also appears to be important for catalysis. Deletion of amino acids 1–160 was shown to significantly reduce enzyme activity [71]. NMR experiments suggested that these 160 N-terminal residues interact with both the remainder of the BPL and the BCCD substrate, although the observed chemical shift perturbations used to measure this were very small [71]. Therefore, residues 1–160 may also be involved in substrate recognition as well as the domain containing residues 160–314. Additional evidence for this comes from investigations into the three naturally occurring human BPL isoforms that arise through differential splicing. Translation of these isoforms is initiated at methionines 1 (full-length BPL), 7, or 58. It has been shown that full-length human BPL has a twofold greater affinity for the BCCD from propionyl-CoA carboxylase than does the Met-58 isoform [60]. However, the biological significance of the three isoforms is not known.

Taken together, the current literature suggests the N-terminal extension characteristic of class III BPLs aids recognition of the BCCD substrate to ensure the selection of appropriate substrates. To allow substrate differentiation as a part of the verification process, BCCDs must contain deviations from the generalized tertiary structure. Comparisons between *E. coli* and human BCCDs indicate subtle structural differences, particularly in the hairpin loop containing the AMKM motif. The *E. coli* BCCD features a type I′ hairpin with the side-chain geometry presenting a clockwise orientation. By contrast, the human acetyl-CoA carboxylase protein has a type I hairpin and counter-clockwise geometry of its side chains [71]. Hence, the specificity of biotinylation in class III organisms appears to be driven by specific interactions between the BCCD and the unique N-terminal extension on the BPL. Further understanding of this species specificity will require the determination of crystal structures of this BPL class and structures of BPL:BCCD complexes.

Concluding Remarks and Future Perspectives

Protein biotinylation represents an important post-translational modification that is necessary for the activation of certain metabolic enzymes. Eukaryotic organisms auxotrophic for biotin must tightly regulate their use of biotin owing to the fine balance of supply and demand. Therefore, these organisms have evolved more complex BPLs through the acquisition of an N-terminal extension. Understanding this class of BPLs has implications for the treatment of MCD. Patients with particular N-terminal domain mutations have no available treatments and the resulting effects can be lethal. Therapeutics that stabilize the interaction between BPL and its protein substrates may provide a treatment solution for patients who are unresponsive to biotin. Protein–protein interaction stabilizers, including peptides and small molecules, are currently being developed as therapeutics in many other disease areas, including as cancer therapeutics and immunosuppressive agents [72]. This approach may also hold promise for MCD patients unresponsive to biotin treatment.

Outstanding Questions

What is the atomic resolution structure of a class III BPL, in particular its N-terminal domain? How conserved is this structure between different eukaryotic BPLs?

What is the function of the N-terminal domain unique to class III BPLs? How does this domain interact with its protein substrate?

What is the molecular basis for the apparent 'substrate verification' activity of the N-terminal domains in class III BPLs?

Can human BPL discriminate between biotin-dependent substrates *in vivo*? Does this selectivity provide a mechanism to ensure the thrifty use of biotin? Are similar mechanisms present in other species with multiple targets for biotinylation?

Three different isoforms of human BPL arise through differential splicing. What are the biological roles of the three isozymes, and do the different isoforms have distinct subcellular localizations and/or substrate selectivity?

How will this new knowledge allow the design of therapeutics to treat MCD patients non-responsive to biotin? In addition, will understanding substrate selectivity change how biotinylation can be used in biotechnological applications by allowing specific labeling?

Is biotin availability and its use linked to other human diseases? For example, the biotin-dependent enzyme pyruvate carboxylase is required for gluconeogenesis, a metabolic pathway often associated with cancer.

Acknowledgments

This work was supported by the National Health and Medical Research Council of Australia (application APP1068885). K.W. is supported by a Ramsay Fellowship from the University of Adelaide. L.S. is the recipient of an Australian Postgraduate Award. We are grateful to the Wallace and Carthew families for their financial support of this work.

References

1. de Oliveira, P.S.L. *et al.* (2016) Revisiting protein kinase–substrate interactions: toward therapeutic development. *Sci. Signal.* 9, re3
2. Beenstock, J. *et al.* (2016) How do protein kinases take a selfie (autophosphorylate)? *Trends Biochem. Sci.* 41, 938–953
3. Patel, D.J. (2016) A structural perspective on readout of epigenetic histone and DNA methylation marks. *Cold Spring Harb. Perspect. Biol.* 8, a018754
4. Falnes, P.O. *et al.* (2016) Protein lysine methylation by seven-β-strand methyltransferases. *Biochem. J.* 473, 1995–2009

Trends in Biochemical Sciences

CellPress

5. Cronan, J.E. (2016) Assembly of lipoic acid on Its cognate enzymes: an extraordinary and essential biosynthetic pathway. *Microbiol. Mol. Biol. Rev.* 80, 429–450
6. Tong, L. (2013) Structure and function of biotin-dependent carboxylases. *Cell. Mol. Life Sci.* 70, 863–891
7. Chapman-Smith, A. and Cronan, J.E.J. (1999) The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem. Sci.* 24, 359–363
8. Lin, S. and Cronan, J.E. (2011) Closing in on complete pathways of biotin biosynthesis. *Mol. Biosyst.* 7, 1811–1821
9. Azhar, A. *et al.* (2015) Mechanisms of biotin transport. *Biochem. Anal. Biochem.* 4, 210
10. Satiaputra, J. *et al.* (2016) Mechanisms of biotin-regulated gene expression in microbes. *Synthet. Sys. Biotechnol.* 1, 17–24
11. Cronan, J.E. (2014) Biotin and lipoic acid: synthesis, attachment and regulation. *EcoSal Plus* 6, 0001–2012
12. Polyak, S.W. *et al.* (2012) Biotin. In *Micronutrients: Sources, Properties and Health Benefits* (Betancourt, A.I. and Gaitan, H. F., eds), pp. 65–93, Nova Science Publishers
13. Wolf, B. (2016) Biotinidase deficiency and our champagne legacy. *Gene* 589, 142–150
14. Said, H.M. (2012) Biotin: biochemical, physiological and clinical aspects. *Subcell. Biochem.* 56, 1–19
15. Ma, Q. *et al.* (2014) Active site conformational changes upon reaction intermediate biotinyl-5′-AMP binding in biotin protein ligase from *Mycobacterium tuberculosis*. *Protein Sci.* 23, 932–939
16. Gupta, V. *et al.* (2010) Structural ordering of disordered ligand-binding loops of biotin protein ligase into active conformations as a consequence of dehydration. *PLoS One* 5, e9222
17. Duckworth, B.P. *et al.* (2011) Bisubstrate adenylation inhibitors of biotin protein ligase from *Mycobacterium tuberculosis*. *Chem. Biol.* 18, 1432–1441
18. Bockman, M.R. *et al.* (2015) Targeting *Mycobacterium tuberculosis* biotin protein ligase (MtBPL) with nucleoside-based bisubstrate adenylation inhibitors. *J. Med. Chem.* 58, 7349–7369
19. Bagautdinov, B. *et al.* (2008) Protein biotinylation visualized by a complex structure of biotin protein ligase with a substrate. *J. Biol. Chem.* 283, 14739–14750
20. Bagautdinov, B. *et al.* (2005) Crystal structures of biotin protein ligase from *Pyrococcus horikoshii* OT3 and its complexes: structural basis of biotin activation. *J. Mol. Biol.* 353, 322–333
21. Tron, C.M. *et al.* (2009) Structural and functional studies of the biotin protein ligase from *Aquifex aeolicus* reveal a critical role for a conserved residue in target specificity. *J. Mol. Biol.* 387, 129–146
22. Weaver, L.H. *et al.* (2001) Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator. *Proc. Natl. Acad. Sci. U. S. A.* 98, 6045–6050
23. Wilson, K.P. *et al.* (1992) *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc. Natl. Acad. Sci. U. S. A.* 89, 9257–9261
24. Wood, Z.A. *et al.* (2006) Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution. *J. Mol. Biol.* 357, 509–523
25. Pendini, N.R. *et al.* (2013) Structural characterization of *Staphylococcus aureus* biotin protein ligase and interaction partners: an antibiotic target. *Protein Sci.* 22, 762–773
26. Beckett, D. (2009) Biotin sensing at the molecular level. *J. Nutr.* 139, 167–170
27. Tissot, G. *et al.* (1997) Evidence for multiple forms of biotin holocarboxylase synthetase in pea (*Pisum sativum*) and in *Arabidopsis thaliana*: subcellular fractionation studies and isolation of a cDNA clone. *Biochem. J.* 323, 179–188
28. Tissot, G. *et al.* (1998) Purification and properties of the chloroplastic form of biotin holocarboxylase synthetase from *Arabidopsis thaliana* overexpressed in *Escherichia coli*. *Eur. J. Biochem.* 258, 586–596
29. Pendini, N.R. *et al.* (2008) Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency. *Biochim. Biophys. Acta* 1784, 973–982
30. Lane, M.D. *et al.* (1964) The enzymatic synthesis of holotranscarboxylase from apotranscarboxylase and (+)-biotin. II. Investigation of the reaction mechanism. *J. Biol. Chem.* 239, 2865–28671
31. Xu, Y. and Beckett, D. (1994) Kinetics of biotinyl-5′-adenylate synthesis catalyzed by the *Escherichia coli* repressor of biotin biosynthesis and the stability of the enzyme-product complex. *Biochemistry* 33, 7354–7360
32. Naganathan, S. and Beckett, D. (2007) Nucleation of an allosteric response via ligand-induced loop folding. *J. Mol. Biol.* 373, 96–111
33. Mayende, L. *et al.* (2012) A novel molecular mechanism to explain biotin-unresponsive holocarboxylase synthetase deficiency. *J. Mol. Med.* 90, 81–88
34. Soares da Costa, T.P. (2012) Selective inhibition of biotin protein ligase from *Staphylococcus aureus*. *J. Biol. Chem.* 287, 17823–17832
35. Purushothaman, S. (2008) Ligand specificity of group 1 biotin protein ligase of *Mycobacterium tuberculosis*. *PLoS One* 3, e2320
36. Kwon, K. *et al.* (2002) Binding specificity and the ligand dissociation process in the *E. coli* biotin holoenzyme synthetase. *Protein Sci.* 11, 558–570
37. Choi-Rhee, E. *et al.* (2004) Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci.* 13, 3043–3050
38. Rees, J.S. *et al.* (2015) Protein neighbors and proximity proteomics. *Mol. Cell. Proteomics* 14, 2848–2856
39. Polyak, S.W. *et al.* (2001) Mutational analysis of protein substrate presentation in the post-translational attachment of biotin to biotin domains. *J. Biol. Chem.* 276, 3037–3045
40. Ingaramo, M. and Beckett, D. (2011) Biotinylation, a post-translational modification controlled by the rate of protein–protein association. *J. Biol. Chem.* 286, 13071–13078
41. Nenortas, E. and Beckett, D. (1996) Purification and characterization of intact and truncated forms of the *Escherichia coli* biotin carboxyl carrier subunit of acetyl-CoA carboxylase. *J. Biol. Chem.* 271, 7559–7567
42. Chapman-Smith, A. (1999) Molecular recognition in a post-translational modification of exceptional specificity: mutants of the biotinylated domain of acetyl-CoA carboxylase defective in recognition by biotin protein ligase. *J. Biol. Chem.* 274, 1449–1457
43. Beckett, D. *et al.* (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* 8, 921–929
44. Solbiati, J. and Cronan, J.E. (2010) The switch regulating transcription of the *Escherichia coli* biotin operon does not require extensive protein–protein interactions. *Chem. Biol.* 17, 11–17
45. Athappilly, F.K. and Hendrickson, W.A. (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* 3, 1407–1419
46. Yao, X. *et al.* (1997) Structure of the carboxy-terminal fragment of the apo-biotin carboxyl carrier subunit of *Escherichia coli* acetyl-CoA carboxylase. *Biochemistry* 36, 15089–15100
47. Roberts, E.L. *et al.* (1999) Solution structures of apo and holo biotinyl domains from acetyl coenzyme A carboxylase of *Escherichia coli* determined by triple-resonance nuclear magnetic resonance spectroscopy. *Biochemistry* 38, 5045–5053
48. Cronan, J.E., Jr (1990) Biotination of proteins *in vivo*. A post-translational modification to label, purify, and study proteins. *J. Biol. Chem.* 265, 10327–10333
49. Stolz, J. *et al.* (1998) Bacteriophage lambda surface display of a bacterial biotin acceptor domain reveals the minimal peptide size required for biotinylation. *FEBS Lett.* 440, 213–217
50. Reche, P. *et al.* (1998) Selectivity of post-translational modifiction in biotinylated proteins: the carboxy carrier protein of the acetyl-CoA carboxylase of *Escherichia coli*. *Biochem. J.* 329, 589–596
51. Brocklehurst, S.M. and Perham, R.N. (1993) Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: a new automated method for the prediction of protein tertiary structure. *Protein Sci.* 2, 626–639

Trends in Biochemical Sciences

CellPress

52. Pares, S. *et al.* (1994) X-ray structure determination at 2.6-A resolution of a lipoate-containing protein: the H-protein of the glycine decarboxylase complex from pea leaves. *Proc. Natl. Acad. Sci. U. S. A.* 91, 4850–4853

53. Reche, P.A. *et al.* (2000) Heteronuclear NMR studies of the specificity of the post-translational modification of biotinyl domains by biotinyl protein ligase. *FEBS Lett.* 479, 93–98

54. Reche, P.A. (2000) Lipoylating and biotinylating enzymes contain a homologous catalytic module. *Protein Sci.* 9, 1922–1929

55. Reche, P. and Perham, R.N. (1999) Structure and selectivity in post-translational modification – attaching the biotinyl-lysine and lipoyl-lysine swinging arms in multifunctional enzymes. *EMBO J.* 18, 2673–2682

56. Cui, G. *et al.* (2006) Identification and solution structures of a single biotin/lipoyl attachment protein from *Bacillus subtilis*. *J. Biol. Chem.* 281, 20598–20607

57. Shenoy, B.C. *et al.* (1992) The importance of methionine residues for the catalysis of the biotin enzyme, transcarboxylase. Analysis by site-directed mutagenesis. *J. Biol. Chem.* 267, 18407–18412

58. Donti, T.R. *et al.* (2016) Holocarboxylase synthetase deficiency pre and post newborn screening. *Mol. Genet. Metab. Rep.* 7, 40–44

59. Campeau, E. and Gravel, R.A. (2001) Expression in *Escherichia coli* of N- and C-terminally deleted human holocarboxylase synthetase. *J. Biol. Chem.* 276, 12310–12316

60. Ingaramo, M. and Beckett, D. (2009) Distinct amino termini of two human HCS isoforms influence biotin acceptor substrate recognition. *J. Biol. Chem.* 284, 30862–30870

61. Healy, S. *et al.* (2010) Structural impact of human and *Escherichia coli* biotin carboxyl carrier proteins on biotin attachment. *Biochemistry* 49, 4687–4694

62. Ingaramo, M. and Beckett, D. (2012) Selectivity in post-translational biotin addition to five human carboxylases. *J. Biol. Chem.* 287, 1813–1822

63. Bailey, L.M. *et al.* (2010) Holocarboxylase synthetase: correlation of protein localisation with biological function. *Arch. Biochem. Biophys.* 496, 45–52

64. Leon-Del-Rio, A. *et al.* (1995) Isolation of a cDNA encoding human holocarboxylase synthetase by functional complementation of a biotin auxotroph of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 92, 4626–4630

65. Polyak, S.W. *et al.* (1999) Biotin protein ligase from *Saccharomyces cerevisiae*: the N-terminal domain is required for complete activity. *J. Biol. Chem.* 274, 32847–32854

66. Kelley, L.A. *et al.* (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Prot.* 10, 845–858

67. Hassan, Y.I. *et al.* (2009) N- and C-terminal domains in human holocarboxylase synthetase participate in substrate recognition. *Mol. Genet. Metab.* 96, 183–188

68. Wilson, C.J. *et al.* (2005) Severe holocarboxylase synthetase deficiency with incomplete biotin responsiveness resulting in antenatal insult in samoan neonates. *J. Pediatr.* 147, 115–118

69. Suormala, T. *et al.* (1997) Five patients with a biotin-responsive defect in holocarboxylase formation: evaluation of responsiveness to biotin therapy *in vivo* and comparative biochemical studies *in vitro*. *Pediatr. Res.* 41, 666–673

70. Santer, R. *et al.* (2003) Partial response to biotin therapy in a patient with holocarboxylase synthetase deficiency: clinical, biochemical, and molecular genetic aspects. *Mol. Genet. Metab.* 79, 160–166

71. Lee, C.-K. *et al.* (2010) The N-terminal domain of human holocarboxylase synthetase facilitates biotinylation via direct interaction with the substrate protein. *FEBS Lett.* 584, 675–680

72. Petta, I. *et al.* (2016) Modulation of protein–protein interactions for the development of novel therapeutics. *Mol. Ther.* 24, 707–718

73. Said, H.M. (2011) Intestinal absorption of water-soluble vitamins in health and disease. *Biochem. J.* 437, 357–372

74. Bandaralage, S.P. (2016) Antenatal and postnatal radiologic diagnosis of holocarboxylase synthetase deficiency: a systematic review. *Pediatr. Radiol.* 46, 357–364

75. Berg, A. *et al.* (1996) Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. *J. Mol. Biol.* 261, 432–442

CHAPTER 2:

FURTHER INTRODUCTION

CHAPTER OUTLINE

Chapter 2 provides additional introductory information necessary for the scope of this thesis and supports the background information summarised in the literature review in chapter 1. Chapter 2 highlights the need for new antifungals to treat human infections, and for the protection of agricultural crops from fungal diseases. One potential new target for the development of novel antifungal therapeutics is biotin protein ligase (BPL), which was extensively reviewed in chapter 1. Here in chapter 2, evidence to support the hypothesis of BPL as a viable antifungal target is provided, including details of the importance of biotin for fungal survival and virulence. Finally, the aims of this project and an outline of the research in this thesis are provided.

2.1 Clinical need for novel antifungal treatments

Antimicrobial resistance has recently gained worldwide attention as a serious threat to humanity, with resistant infections predicted to be responsible for over 10 million deaths by 2050 (1). Whilst most of this attention has been focused on bacterial resistance to current antibiotics, the threat of fungal infections and emerging resistance to antifungal treatments is under-appreciated (2). The most prevalent fungal infections occur superficially on the skin and nails, or at mucosal surfaces. However, severe invasive tissue or systemic bloodstream infections are an under-recognised threat (2,3). Invasive, systemic fungal infections affect over a billion people annually and result in high mortality rates of over 1.5 million deaths per year (2,3), despite the availability and administration of antifungal therapeutics (4). The four fungal genera responsible for causing 90% of deaths from invasive infections are *Cryptococcus, Candida, Aspergillus* and *Pneumocystis* (2). Within each of these genera there are several species that can cause disease, however, often one or a few specific species are the most prevalent. For example *Cryptococcus neoformans* and *Cryptococcus gattii* are responsible for most cases of cryptococcosis, whilst *Candida abicans, Aspergillus fumigatus, Aspergillus flavus* and *Pneumocystis jirovecil* are the most common species of *Candida, Aspergillus* and *Pneumocystis* responsible for causing disease (2).

Antifungal chemotherapies are available for invasive fungal infections, however, these do not always prevent mortality (4). Whilst the total number of fungal treatments on the market has increased during the past 15 years, only one new class of antifungals has been licenced during this time, namely the echinocandins (5,6). Alternative antifungal classes currently used to treat fungal infections include the azoles, polyenes and nucleoside analogues (7,8). Azoles, one of the most frequently prescribed antifungal classes, inhibit ergosterol synthesis causing the production of toxic bi-products from this pathway and reducing the amount of ergosterol in the cell membrane leading to the loss of membrane structure and function (Figure 1) (9). Polyenes also act upon sterols, binding to these molecules in the cell membrane to form pores that disrupt the osmotic balance of the cell. Echinocandins inhibit cell wall synthesis by inhibiting β-D-glucan synthase, which prevents the production of the important cell wall component β-1,3-D glucan (Figure 1) (10). The azoles and echinocandins are the current first-line antifungals for the treatment of *Candida* infections, however, resistance has and continues to be detected (7,9,11). *Aspergillus* species resistant to azoles have also been identified, whilst *Cryptococcus* species are not susceptible to the echinocandins (10,12). Resistance to polyenes has also been detected for *Aspergillus*

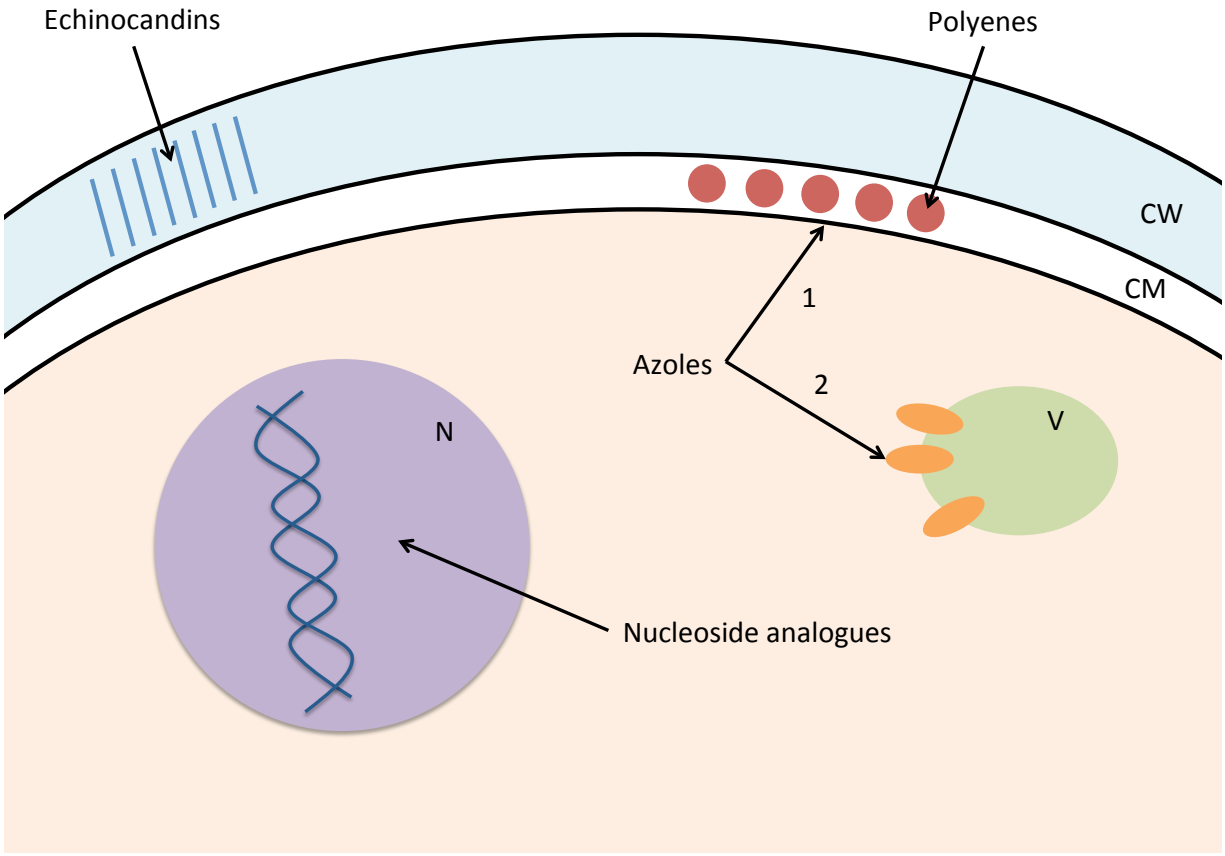


Figure 1. Mechanisms of action of current classes of antifungal drugs. The echinocandins inhibit β(1-3)D-glucan synthesis (blue rectangles), a component of the cell wall (CW), reducing the integrity of the cell wall. The azoles inhibit ergosterol (red circle) biosynthesis (arrow 1) through the target lanosterol demethylase (ERG11). Decreased cell membrane (CM) sterol production leads to the accumulation of toxic bi-products and membrane stress. Some azoles have also been shown to inhibit vacuolar H^+-ATPases (orange ovals), leading to imbalances in ionic homeostasis (arrow 2). The azoles utilised in agriculture to control plant fungal pathogens also act by the same mechanisms. Polyenes bind ergosterol (red circle) in the membrane and these lipid:drug complexes form membrane-spanning channels. These channels disrupt the ionic balance of the cell and the integrity of the membrane. Finally, the nucleoside analogues inhibit DNA replication, RNA synthesis and protein synthesis (10). CW: cell wall, CM: cell membrane, N: nucleus, V: vacuole (green oval).

species, and also at relatively low incidence rates for *Candida* (12,13). The nucleoside analogues function via a distinct mechanism, mimicking naturally occurring nucleosides such that they can be incorporated into DNA or RNA to disrupt DNA, RNA and protein synthesis (Figure 1) (8,10). *Candida* resistance to flucytosine, a specific nucleoside analogue, has been reported but only at low incidence rates (8,12). Resistant infections can still be treated using antifungals from different classes that act through different mechanisms. However, this can be associated with significant treatment costs and cytotoxic side effects, with the polyenes and first generation azoles causing nephrotoxicity and liver damage (14). Side effects occur in 50-90% of cases when amphotericin B, a specific polyene, is administered (15). Likewise, the availability of appropriate treatments and their formulations impede the treatment of resistant infections in some developing countries, as they may have limited access to only a single antifungal class (7). Hence, infections resistant to multiple classes of antifungals can be left with few or even no treatment options (7). For example, strains of *Candida auris* have emerged that possess high resistance levels to multiple classes of antifungals (16), with some isolates resistant to all antifungal classes used to treat *Candida* infections. Therefore, there are no remaining treatment options for this species (17).

The resistance and toxicity associated with current antifungals, together with the high mortality rates of 35-55% for fungal infections even when antifungals are administered, demonstrates the need for new treatments that work through novel mechanisms of action to treat systemic, invasive fungal infections. As fungi are evolutionarily close to humans, there are few fungal targets that lack a corresponding human homologue. Hence, species selectivity must be considered when developing therapeutics, to prevent potential toxicity (2,18).

2.2 Agricultural need for new antifungal therapeutics

Fungal infections are not only a serious threat to human health but are also responsible for various diseases in plants (19). Fungi, yeast and moulds infect many commercially valuable crops including wheat, fruits, wine grapes, flowers and lentils, reducing the yield and quality of the produce (19-22). This has a detrimental effect on the agricultural industry, reducing crop yields and profits, and impacting the security of the world's food supply (19). Billions of dollars are spent annually on fungicide treatment, but resistance to these agents has been detected and continues to emerge (19,23). Two of the most problematic global fungal pathogens include *Botrytis cinerea* and *Zymoseptoria tritici* (19).

Botrytis cinerea causes grey mould disease (also known as *Botrytis* bunch rot or *Botrytis* fruit rot) and affects over 1000 species of plants worldwide (24). It infects the leaves, stems, flowers and fruits of many economically important crops including berries (especially strawberries), tomatoes, ornamental flowers (including roses, petunias and gerberas), wine and table grapes, and lentils (19-21,25). Not only can *B. cinerea* infect plants during growth from seedling to ripening in open fields or greenhouses, but it can also cause spoilage and damage post-harvest to seemingly healthy crops during storage, transport and retail, even if the produce are kept refrigerated (0 to 10°C) (19,20). *B. cinerea* is a typical necrotrophic mould with complex penetration, infection and colonisation strategies that help it evade the host's defences and promote programmed cell death of the host plant tissue (20,24,26). *B. cinerea* is opportunistic, infecting weak, damaged, mature or senescent tissues. It can also remain quiescent for considerable amounts of time before causing damage when the host's physiology changes or the environmental conditions become more favourable (19,26).

B. cinerea is a major plant pathogen that infects a wide range of hosts at multiple stages of growth and/or production (19). It is estimated that 15% to 40% of berries, tomatoes and ornamental flowers are lost to post-harvest spoilage (20), but these predictions may be vastly under-estimated as it is difficult to quantify the costs associated with losses at each stage of the production process for a broad range of crops (19). *B. cinerea* is also a serious threat to Australian lentil production, with 2001 yield losses of 15-25% in South Australia and 2-10% in Victoria. At an average yield spoilage of 11% a year, this results in the loss of $13 million (21). Similarly, the flower industry is impacted by *B. cinerea* reducing flower quality and vase-life, resulting in revenue losses of millions (19). Overall, the economic damage from *B. cinerea* has been estimated in excess of €1 billion annually (~$1.57 billion AUD at time of

writing) (24). The primary method for controlling *B. cinerea* is through the application of chemical fungicides (20,25,26). Specific fungicides for use on *Botrytis* are called botryticides and compose 10% of the global fungicide market (19). Broad-spectrum fungicides are also used, with the average cost of chemical treatment estimated at €40 per hectare (~$60 AUD at time of writing) (19). This widespread, extensive fungicide use has resulted in the emergence of resistance, including species that demonstrate multifungicide resistance (19,25).

Another important fungal pathogen in the agricultural industry is *Zymoseptoria tritici*, also known as *Mycosphaerella graminicola* (27). This fungus causes the disease *Septoria tritici* blotch that affects many wheat crops along with oats and barley (22). *Z. tritici* is the most damaging wheat disease in Europe and a serious threat to many other global wheat production areas (27). Wheat is a precious resource as it is the second most important food crop required for human and animal feed sources, and is the raw material for biofuels and many other resources (28). *Z. tritici* infections cause brown coloured lesions that run along the veins of the leaf and small, black spots, termed pycnidia, that form on the dead leaf tissue. These spots produce spores that spread the infection to new plant hosts via rain splashes or surviving in wheat stubble across seasons (22,29). *Z. tritici* penetrates into the host plant cells approximately 12 hours post inoculation (30), where it resides asymptomatically and evades host recognition for long periods of time (from 7 to 28 days) (19,28,29,31). The pathogen then converts to an aggressive necrotrophic lifestyle, promoting host programmed cell death and causing symptoms to appear (19,30). During this aggressive lifestyle, *Z. tritici* degrades and utilises host macromolecules as a source of nutrients for its own growth and sporulation (30).

In extreme circumstances, *Z. tritici* causes yield losses of up to 50%, with yield losses of 20% typical across the UK (23,27,32). Losses of 5-10% are more common when fungicides or wheat varieties that are resistant to *Z. tritici* infection are utilised (23). These losses are estimated to cost €120-700 million (~$180 million to $1.1 billion AUD at time of writing) across the major European wheat producers of France, Germany and the UK (28). The management of *Z. tritici* makes up 70% of the annual cereal fungicide usage in Europe, costing up to €1 billion annually (~$1.57 billion AUD at time of writing) (23,27,28,30). Likewise, fungicide management of *Z. tritici* in the USA is also expensive, costing US$275 million anually (~$380 million AUD at time of writing) (30). Despite the availability of fungicides and the yield improvements they facilitate, *Z. tritici* is developing resistance to

these treatments. Resistance levels to specific antifungals vary globally. Widespread resistance to the quinone outside inhibitor fungicide class is present in the UK such that these chemicals are no longer effective (19,22,23). Australia is one of the few places where these antifungals remain useful due to low rates of resistance (22,33), despite large-scale resistance to other fungicide classes in Australia (22,30,33).

These two fungal pathogens, along with many others, demonstrate the urgent need for new antifungals for use in the agricultural industry to control fungal infections in crops. New fungicides are required to combat current and emerging resistance, and to complement the few existing fungicide classes (23). Control of pathogens in agriculture is crucial to secure society's food resources, and to allow an increase in food production that will be required to support global population growth (30,34).

2.3 Biotin physiology in fungi

Fungi are auxotrophic for biotin, acquiring this essential micronutrient from the environment for 2 major roles: as 1) an enzyme co-factor and 2) a morphological regulator, the latter in the the fungus *Candida albicans* in particular (35). Most of the research into the role of biotin and its related pathways in fungi has been carried out in the model yeast *S. cerevisiae* and the human pathogen *C. albicans.* Hence, these two fungal species will be discussed to provide an insight into the biotin physiology of fungi. *C. albicans* can exist in multiple different morphologies, including unicellular, budding yeast and an elongated hyphal, filamentous form (36). The fungi can reversibly switch its morphology depending on the external environment, and can be present in either the yeast or filamentous forms in the host during infection (36,37). The morphology of *C. albicans* is correlated with its pathogenicity, as virulence factors are co-regulated with the genes for hyphal morphogenesis (11,37). Hence, the hyphal morphology is a crucial attribute to *C. albicans* virulence (37). Biotin can induce hyphal morphogenesis, as increased biotin concentrations in the growth media enhanced *Candida* hyphal formation in a dose-dependent manner (35). Therefore, biotin has an important signalling role in inducing the hyphal morphology of *C. albicans* that is associated with promoting *Candida* infections. The precise signalling mechanisms and pathways by which biotin induces morphological changes in *Candida* have not been determined.

In its cofactor role, biotin is covalently attached to biotin-dependent enzymes where it facilitates their activity to catalyse essential metabolic processes. The biotin moiety is responsible for carrying CO_2 groups during the carboxylation, decarboxylation and transcarboxylation reactions of the biotin-dependent enzymes (see Chapter 1) (38). Five biotin-dependent enzymes have been identified in *Saccharomyces cerevisae*. Genomic analysis reveals a similar number for other fungi including *C. albicans*. These biotin-dependent enzymes include two isoforms of acetyl CoA carboxylase, two isoforms of pyruvate carboxylase, and a urea amidolyase (Figure 2) (39). The cytoplasmic isoform of acetyl CoA carboxylase completes the rate-limiting step of lipogenesis to produce malonyl CoA, which is utilised for the elongation of fatty acid chains, and the synthesis and maintenance of membranes (40). Disruption of the *acc1* gene is lethal to *S. cerevisiae*, even when supplemented with fatty acids for growth (40). Similarly, the mitochondrial isoform of acetyl CoA carboxylase, HFA, is also responsible for fatty acid synthesis and has also been implicated in lipoic acid synthesis (41). It is responsible for producing malonyl-CoA for utilisation in the mitochondria, as this fatty acid precursor is impervious to the mitochondrial

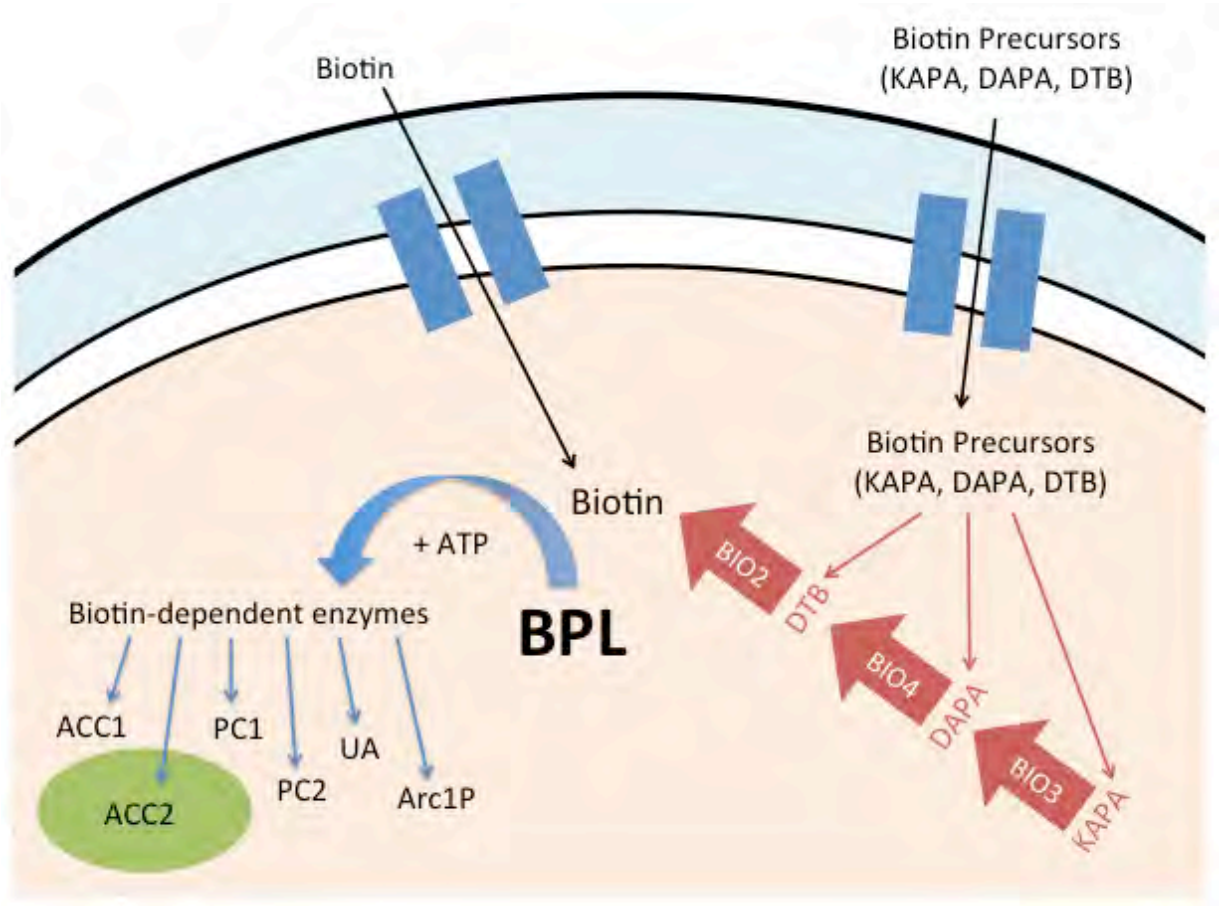


Figure 2. Predicted biotin biology pathways within fungal cells, based on empirical data from *S. cerevisiae* and *C. albicans.* Biotin is transported into the cell via a high affinity, energy-requiring transporter (VHT1). Here, BPL covalently attaches biotin to biotin-dependent enzymes via an ATP-dependent reaction. There are five biotin-dependent enzymes, namely acetyl-coA carboxylase isoform 1 (ACC1), acetyl-coA carboxylase isoform 2 (ACC2) (present in the mitochondria (shown in green)), pyruvate carboxylase isoform 1 (PC1), pyruvate carboxylase isoform 2 (PC2) and urea amidolyase (UA), although the protein Arc1p can also be biotinylated. Fungi can also utilise the biotin precursors KAPA, DAPA and DTB, which can be transported into and accumulated in the cell. BIO5 has been identified as a transporter for KAPA and VHT1 is proposed to also facilitate DTB uptake, however, a DAPA transporter has not been identified. The partial biosynthesis pathway (in red) comprising DAPA synthase (BIO3), dethiobiotin synthetase (BIO4) and biotin synthase (BIO2), can utilise these precursors to synthesise biotin.

membrane (41). HFA mutants are viable, but contain no lipoic acid and are respiratory defective, only growing on fermentative media (42). The two isoforms of pyruvate carboxylase, encoded by separate genes, play a key role in gluconeogenesis, catalysing the anaplerotic reaction to regenerate oxaloacetate stores for the citric acid cycle and energy production (43,44). Disruption of either isoform did not result in a marked change to the *S. cerevisiae* phenotype, suggesting some redundancy between the isoforms (45). However, double knockouts of both isoforms are unable to grow when glucose is the sole carbon source unless the growth media is supplemented with aspartate (45). Urea amidolyase is an energy-requiring cytoplasmic enzyme that completes the first strep of urea metabolism in order to use urea as a sole nitrogen source (46,47). Urea amidolyase knockout mutants of *C. albicans* were unable to use nitrogen as an energy source and showed a reduction in survival and virulence. This suggested urea amidolyase may be a *C. albicans* virulence factor, as it often is in other bacteria and fungi (47). Another protein has been identified to be biotinylated in *S. cerevisiae*, although this protein is not an enzyme but rather a tRNA binding-protein called Arc1p (48). Although Arc1p is biotinylated specifically by BPL it does not carry out biotin-dependent carboxylation, decarboxylation or transcarboxylation reactions (48). Recent literature revealed only low levels of Arc1p biotinylation in *S. cerevisiae* and that biotinylation is not required for Arc1p activity with the exact role of this modification remaining unknown (48,49). A homologue of Arc1p was also identified in *C. albicans*, however, it is not biotinylated *in vivo* in contrast to the *S. cerevisiae* protein (49). Despite this, the 5 other biotin-dependent enzymes demonstrate that biotin is involved in many vital metabolic processes for the survival and virulence of fungal organisms.

Most mammals and fungi are auxotrophic for biotin, obtaining this micronutrient from the environment via their diet (38). Acquisition of biotin from environmental sources occurs through high affinity transport channels (reviewed (50)). *S. cerevisiae* obtains biotin by importing the micronutrient through its energy-dependent, high affinity Vitamin H-Transporter 1 (VHT1) (51). This channel has a high affinity for biotin, with the K_M reported as 3.23×10^{-7} M (52). The *C. albicans* biotin transporter has not been characterised, but there is a predicted ortholog of VHT1 at *orf19.2397* that is hypothesised to act in a similar manner to VHT1 (53). Through its biotin transporter, *S. cerevisiae* can accumulate biotin to a concentration up to 1000 times the external biotin concentration (51,52,54). As external biotin concentrations increase, there is a decrease in the specific uptake rate by *S. cerevisiae* VHT1 and an efflux of any extra-accumulated biotin from the cell (52,54,55). Therefore, transport of biotin is finely

tuned to respond to internal and external biotin concentrations through mechanisms not yet determined in fungi.

Although fungi are auxotrophic for biotin and contain high affinity transporters for its uptake, some fungi also contain partial biotin synthesis pathways (56-58). The pathway responsible for biosynthesis is conserved in select organisms capable of making their own biotin, such as bacteria and plants. This pathway consists of various enzymes that convert a malonyl-CoA precursor to *d*-biotin (Figure 3) (reviewed in (59)). Whilst *C. albicans* and *S. cerevisiae* cannot produce biotin *de novo,* they have retained a partial biotin synthesis pathway comprised of the final three enzymes of the pathway (35,56,57). The acquisition of this partial pathway in fungi is predicted to occur from the loss of the complete synthesis pathway from a common fungal ancestor, and then the rebuilding of the pathway by horizontal gene transfer and gene duplication, together with neofunctionalization (when a gene acquires a new function after gene duplication that was not present in the ancestral gene) (58). Therefore, *S. cerevisiae* and *C. albicans* can fulfil their auxotrophic requirement for biotin by scavenging the precursors of the last 3 synthetic enzymes namely KAPA, DAPA and dethiobiotin (35,56,60,61). They then utilise the enzymes BIO3 (DAPA synthase), BIO4 (dethiobiotin synthetase) and BIO2 (biotin synthase) to produce biotin from these precursors (Figures 2 and 3) (35,56,57). This can allow the fungi to continue obtaining biotin even when it is not freely available from the environment.

In several bacterial species, the biotin synthesis enzymes are also important for promoting virulence. The second enzyme in this pathway, DAPA synthase (BioA) that converts KAPA to DAPA (Figure 3), was shown to be crucial for viability and virulence in *M. tuberculosis* (62). *M. tuberculosis* BioA knockout mutants demonstrated decreased levels of biotinylated proteins, reduced viability *in vitro*, and were unable to establish and maintain infections in mice (62). The necessity of biotin synthesis, particularly BioA, for the viability and virulence of *M. tuberculosis* has allowed these enzymes to be targeted for novel anti-tuberculosis therapeutics (63-69). Likewise, fungal biotin biosynthesis enzymes have also been shown to be important during infection with *C. albicans* up-regulating its biotin biosynthesis genes during both infection and biotin starvation. BIO2, BIO3 and BIO4 are each up-regulated 11-, 1500- and 150-fold respectively in response to biotin starvation (35). Similar studies reveal BIO2 and BIO4 were up-regulated approximately 2.5-fold during infection (70). These responses suggest biotin is essential for fungal survival and infection, enabling fungi to increase biotin production when necessary to support fungal viability and/or virulence.

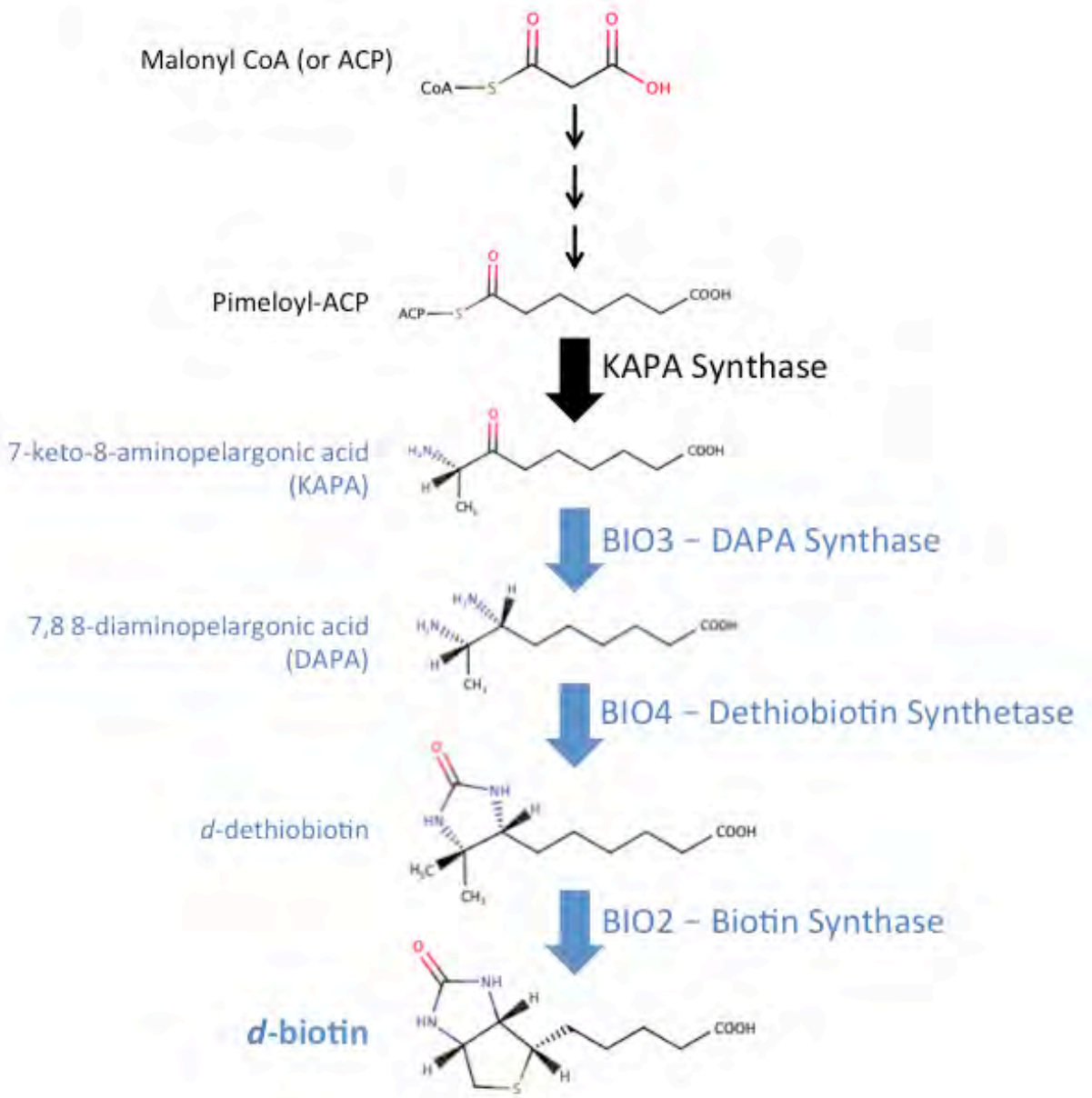


Figure 3. The biotin biosynthetic pathway is conserved in all organisms that produce their own biotin. Malonyl CoA is converted to pimeloyl-ACP by a sequence of enzymes, before pimeloyl enters the conserved biosynthesis pathway. Here it is converted to KAPA by KAPA synthase. DAPA synthase then catalyses the formation of DAPA from KAPA, before dethiobiotin synthetase catalyses the further conversion to dethiobiotin. Finally, biotin synthase converts dethiobiotin to biotin. In *S. cerevisiae* and *C. albicans* only the final three steps of this synthesis pathway are conserved (shown in blue), allowing fungi to produce biotin from KAPA, DAPA or dethiobiotin.

2.4 Biotin Protein Ligase (BPL) as a novel antifungal target

The necessity of biotin for the survival and pathogenicity of fungi suggests that the enzyme that controls biotin usage and attachment, biotin protein ligase (BPL), is a promising antifungal target. Inhibition of BPL prevents the covalent attachment of biotin onto all five biotin-dependent enzymes, thereby reducing their activity and interfering with multiple important metabolic processes including lipogenesis for the synthesis and maintainence of cell membranes, energy production, and the ability to utilise urea as a nitrogen source (40,41,43,44,47). Specifically, it would inhibit metabolic pathways important for fungal viability (acetyl-CoA carboxylase isoforms 1 and 2, and pyruvate carboxylase isoforms 1 and 2) and virulence (urea amidolyase) (40,42,45,47). BPL is also essential for yeast viability. In *S. cerevisiae* genetic knockout studies only half the haploid spores were viable post sporulation and those that grew all possessed a functional *bpl* allele (71). BPL was also validated as an essential enzyme in *S. cerevisiae* by the yeast genome deletion project. Here, each of the ~6000 open reading frames identified in *S. cerevisiae* were systemically deleted to identify those necessary for the viability of the yeast. BPL1, ACC1 and VHT1 were all shown to be essential (72). BPL was also identified as an essential gene for the viability of the fungi *C. albicans* and *Schizosaccharomyces pombe* (73,74). Therefore, BPL is hypothesised to be an attractive novel antifungal target due to its necessity for fungal survival.

Further supporting the hypothesis that BPL is a promising novel antifungal target is the overlap of its proposed mechanism of action to current licenced antifungals. The mechanism of action of the current azole antifungals (used to treat human and agricultural infections) and the demethylation inhibitors (used in agriculture) involves inhibiting sterol synthesis. Sterol molecules are important for maintaining the cell membrane integrity (10,23). Other agricultural antifungals, including the quinone outside inhibitors and succinate dehydrogenase inhibitors, are responsible for inhibiting energy producing pathways including the TCA cycle and electron transport in the respiratory chain (23,75,76). Targeting BPL would affect similar pathways as these licenced drugs. Inhibiting acetyl CoA carboxylase would prevent the synthesis of lipids that are crucial for membrane synthesis and maintenance, and inhibiting pyruvate carboxylase and urea amidolyase interferes with energy production processes (40,41,43,44,47). This further validates BPL as a good antifungal target, as previously inhibiting these pathways has produced licenced antifungals. Resistance to these current antifungals and their mechanisms of action are emerging. Despite this, targeting BPL should still be effective as the actual targets (BPL and the biotin-dependent enzymes) are different and would circumvent most previously existing resistance mechanisms. Furthermore,

inhibiting BPL would interfere with multiple essential metabolic pathways. More than one of these pathways would be required to develop resistance mechanisms in order for the fungi to be able to survive in the presence of a BPL inhibitor. Studies of the spontaneous development of resistance of *S. aureus* and *M. tuberculosis* to BPL inhibitors occurred at rates of less than 1×10^{-9} and 1.4×10^{-10} respectively (77,78). Therefore, resistance to BPL inhibitors is possible but the low incidence rates further strengthen the hypothesis of BPL as an antimicrobial target.

BPL is essential for most, if not all, living organisms. Specifically, several genetic knockout studies have also revealed the essentiality of BPL for bacteria (79-81), which has led to BPL being a target for the development of novel antibiotics, for example against *S. aureus* (82-87) and *M. tuberculosis* (77,81,88,89). Despite the presence of a BPL homologue in humans, selective inhibition of the bacterial *S. aureus* BPL over the human isoform has been achieved (>1100-fold selectivity) (83). This is due to divergence in the active site residues of the bacterial enzyme compared to the human isoform, such that the inhibitor forms specific interactions with the active site of the bacterial enzyme but not the human enzyme (83). It is proposed that a similar mechanism of selectivity could be achieved between fungal pathogenic BPLs and the human isoform, depending on the homology of their active sites. The level of this homology is not precisely known, due to the limited structural data available. Archaeal and bacterial BPLs have been well characterised structurally and biochemically. In contrast, fungal BPLs, which together with other eukaryotic BPLs comprise a unique structural class termed the class III BPLs (reviewed Chapter 1), have not been extensively characterised. In addition to the catalytic biotinylation domain and C-terminal cap, these enzymes contain an N-terminal extension that is predicted to aid recognition of substrates for biotinylation. Currently there is very limited structural information about this BPL class, and no atomic-resolution data. However, such structural knowledge will be vital for pursuing fungal BPLs as novel antifungal targets.

2.5 Previous attempts at structural characterisation of class III BPLs

In order to aid antifungal drug design efforts and gain a thorough understanding of a eukaryotic BPL, our lab has previously attempted to determine the structure of a class III BPL. The approaches utilised have included X-ray crystallography of the full-length *Sc*BPL in the presence and absence of ligands biotin and MgATP. Nuclear magnetic resonance (NMR) was also carried out on the N-terminal structured domain truncation protein from the human BPL (residues 159-293) in an attempt to provide the structure of this novel region (90). It is noteworthy that the size of a full-length class III BPL is too large to be easily amenable to structural determination via NMR (molecular weight of 75-80 kDa compared to the NMR size restraint of approximately 30 kDa). Neither of these approaches were successful at producing high-resolution structures, as no diffraction-quality crystals were obtained, nor soluble, non-aggregating N-terminal extension protein for NMR (90). Therefore, alternative approaches to gain insight into the structure of a class III BPL must be considered due to the difficulty in generating an atomic resolution structure by conventional structural techniques.

2.6 Aims of the project

A structure of a class III BPL remain elusive despite their intimate involvement in human health and disease (multiple carboxylase deficiency, reviewed chapter 1), and their potential as novel antifungal targets. This thesis aims to biochemically and structurally characterise several class III BPLs from the prototypical yeast *S. cerevisiae* and two fungal species of agricultural importance. Here techniques that have not previously been applied to BPLs were explored, including various mass spectrometry approaches. In order to establish these techniques and develop appropriate protocols, the well-studied bacterial *S. aureus* BPL (*Sa*BPL) was utilised. Chapter 3 discusses the establishment of native nano-electrospray ionisation-mass spectrometry (nESI-MS) as a technique that was then applied to analyse the monomer-dimer equilibrium and DNA-binding activity of *Sa*BPL. This technique was also applied to study the self-oligomerisation of a mutant *Sa*BPL, to determine if it homodimerizes similarly to wild-type enzyme (Appendix 2).

Following the establishment of mass spectrometry protocols for *Sa*BPL, these techniques were then employed in conjunction with others to characterise other class III BPLs. The BPLs from *S. cerevisiae, B. cinerea* and *Z. triciti* were purified for characterisation. Structural and kinetic characterisation of these enzymes was completed (Chapter 4) to compare the three different fungal BPLs and assess their potential as targets for novel antifungal therapeutics.

High-resolution structural data are required to aid the antifungal drug development process. Therefore, Chapter 5 focuses on further structural investigation of the *S. cerevisiae* BPL (*Sc*BPL) as a model for other class III BPLs that are involved in pathogenic or agricultural infections. In conjunction with investigating the overall structure of a class III BPL, the conformational changes associated with ligand binding were also studied. Ultimately, structural information about BPL will provide vital insights for the development of antifungals.

The specific aims for this project are:

1. Develop native mass spectrometry protocols to be utilised for structural studies of BPL;
2. Produce and kinetically/structurally characterise the BPLs from different fungal pathogens;
3. Investigate the structure of class III BPLs and the dynamic conformational changes that accompany ligand binding.

2.7 References

1. The Review on Antimicrobial Resistance chaired by Jim O'Neill. (2016) Tackling drug-resistance infections globally: final reports and recommendations.
2. Brown, G. D., Denning, D. W., Gow, N. A. R., Levitz, S. M., Netea, M. G., and White, T. C. (2012) Hidden Killers: Human Fungal Infections. *Sci Transl Med* **4**
3. Bongomin, F., Gago, S., Oladele, R. O., and Denning, D. W. (2017) Global and Multi-National Prevalence of Fungal Diseases - Estimate Precision. *J Fungi* **3**
4. Kullberg, B. J., and Arendrup, M. C. (2015) Invasive Candidiasis. *N Engl J Med* **373**, 1445-1456
5. Denning, D. W. (2002) Echinocandins: a new class of antifungal *J Antimicrob Chemother* **49**, 889-891
6. Stan, C. D., Tuchiluş, C., and Stan, C. I. (2014) Echinocandins--new antifungal agents. *Rev Med Chir Soc Med Nat Iasi* **118**, 528-536
7. World Health Organisation. (2014) Antimicrobial Resistance: Global Report on Surveillance.
8. Spampinato, C., and Leonardi, D. (2013) *Candida* Infections, Causes, Targets, and Resistance Mechanisms: Traditional and Alternative Antifungal Agents. *Biomed Res Int*
9. Sanguinetti, M., Posteraro, B., and Lass-Flörl, C. (2015) Antifungal drug resistance among *Candida* species: mechanisms and clinical impact. *Mycoses* **58**, 2-13
10. Prasad, R., Banerjee, A., and Shah, A. H. (2017) Resistance to antifungal therapies. *Essays Biochem* **61**, 157-166
11. Pappas, P. G., Lionakis, M. S., Arendrup, M. C., Ostrosky-Zeichner, L., and Kullberg, B. J. (2018) Invasive candidiasis. *Nat Rev Dis Primers* **4**
12. Kanafani, Z. A., and Perfect, J. R. (2008) Resistance to Antifungal Agents: Mechanisms and Clinical Impact. *Clin Infect Dis* **46**, 120-128
13. Arendrup, M. C., and Patterson, T. F. (2017) Multidrug-Resistant *Candida*: Epidemiology, Molecular Mechanisms, and Treatment. *J Infect Dis* **216**, S445-S451
14. Campoy, S., and Adrio, J. L. (2017) Antifungals. *Biochem Pharmacol* **133**, 86-96
15. Enoch, D. A., Ludlam, H. A., and Brown, N. M. (2006) Invasive fungal infections: a reveiw of epidemiology and management options. *J Med Microbiol* **55**, 809-818
16. Sears, D., and Schwartz, B. S. (2017) *Candida auris*: An emerging multidrug-resistant pathogen. *Int J Infect Dis* **63**, 95-98
17. Chowdhary, A., Voss, A., and Meis, J. F. (2016) Multi-drug resistant *Candida auris*: 'new kid on the block' in hospital-associated infections? *J Hosp Infect* **94**, 209-212
18. Pianalto, K. M., and Alspaugh, J. A. (2016) New Horizons in Antifungal Therapy. *J Fungi* **2**
19. Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., Rudd, J. J., Dickman, M., Kahmann, R., Ellis, J., and Foster, G. D. (2012) The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol* **13**, 414-430
20. AbuQamar, S., Moustafa, K., and Tran, L. S. (2017) Mechanisms and strategies of plant defense against *Botrytis cinerea*. *Crit Rev Biotechnol* **37**, 262-274
21. Grains Research and Development Corporation. (2008) UM00015 - Epidemiology and control of Botrytis grey mould of lentils.
22. Grains Research & Development Corporation. (2014) Septoria Tritici Blotch Fact Sheet.
23. Torriani, S. F. F., Melichar, J. P. E., Mills, C., Pain, N., Sierotzki, H., and Courbot, M. (2015) *Zymoseptoria tritici*: A major threat to wheat production, integrated approaches to control. *Fungal Genet Biol* **79**, 8-12

24. Veloso, J., and van Kan, J. A. L. (2018) Many Shades of Grey in *Botrytis*-Host Plant Interactions. *Trends Plant Sci* **23**, 613-622
25. Imada, K., Tanaka, S., Ibaraki, Y., Yoshimura, K., and Ito, S. (2014) Antifungal effect of 405-nm light on *Botrytis cinerea*. *Lett Appl Microbiol* **59**, 670-676
26. AbuQamar, S. F., Moustafa, K., and Tran, L.-S. P. (2016) 'Omics' and Plant Responses to *Botrytis cinerea*. *Front Plant Sci* **7**
27. McDonald, B. A., and Mundt, C. C. (2016) How Knowledge of Pathogen Population Biology Informs Management of Septoria Tritici Blotch. *Phytopathology* **106**, 948-955
28. Fones, H., and Gurr, S. (2015) The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal Genet Biol* **79**, 3-7
29. Steinberg, G. (2015) Cell biology of *Zymoseptoria tritici*: Pathogen cell organization and wheat infection. *Fungal Genet Biol* **79**, 17-23
30. O'Driscoll, A., Kildea, S., Doohan, F., Spink, J., and Mullins, E. (2014) The wheat-*Septoria* conflict: a new front opening up? *Trends Plant Sci* **19**, 602-610
31. McDonald, M. C., McDonald, B. A., and Solomon, P. S. (2015) Recent advances in the *Zymoseptoria tritici*-wheat interaction: insights from pathogenomics. *Front Plant Sci* **6**
32. Milgate, A. (2016) DAN00203 - Effective genetic control of Septoria tritici blotch (STB).
33. Milgate, A. (2017) Septoria tritici blotch rears its head in central and southern NSW (Northern).
34. McDonald, B. A., and Stukenbrock, E. H. (2016) Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. *Philos Trans R Soc Lond B Biol Sci* **371**
35. Ahmad Hussin, N., Pathirana, R. U., Hasim, S., Tati, S., Scheib-Owens, J. A., and Nickerson, K. W. (2016) Biotin Auxotrophy and Biotin Enhanced Germ Tube Formation in *Candida albicans*. *Microorganisms* **4**
36. Lu, Y., Su, C., and Liu, H. (2014) *Candida albicans* hyphal initiation and elongation. *Trends Microbiol* **22**, 707-714
37. Su, C., Yu, J., and Lu, Y. (2018) Hyphal development in *Candida albicans* from different cell states. *Curr Genet* **64**, 1239-1243
38. Sternicki, L. M., Wegener, K. L., Bruning, J. B., Booker, G. W., and Polyak, S. W. (2017) Mechanisms Governing Precise Protein Biotinylation. *Trends Biochem Sci* **42**, 383-394
39. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C. J., Wallace, J. C., and Polyak, S. W. (2008) Biotin protein ligase from *Candida albicans*: Expression, purification and development of a novel assay. *Arch Biochem Biophys* **479**, 163-169
40. Haßlacher, M., Ivessa, A. S., Paltauf, F., and Kohlwein, S. D. (1993) Acetyl-CoA Carboxylase from Yeast Is an Essential Enzyme and Is Regulated by Factors That Control Phospholipid Metabolism. *J Biol Chem* **268**, 10946-10952
41. Hoja, U., Marthol, S., Hofmann, J., Stegner, S., Schulz, R., Meier, S., Greiner, E., and Schweizer, E. (2004) HFA1 Encoding an Organelle-specific Acetyl-CoA Carboxylase Controls Mitochondrial Fatty Acid Synthesis in *Saccharomyces cerevisiae*. *J Biol Chem* **279**, 21779-21786
42. Brody, S., Oh, C., Hoja, U., and Schweizer, E. (1997) Mitochondrial acyl carrier protein is involved in lipoic acid synthesis in *Saccharomyces cerevisiae*. *FEBS Lett* **408**, 217-220
43. Walker, M. E., Val, D. L., Rohde, M., Devenish, R. J., and Wallace, J. C. (1991) Yeast pyruvate carboxylase: identification of two genes encoding isoenzymes. *Biochem Biophys Res Commun* **176**, 1210-1217
44. Brewster, N. K., Val, D. L., Walker, M. E., and Wallace, J. C. (1994) Regulation of Pyruvate Carboxylase Isozyme (*PYC1, PYC2*) Gene Expression in *Saccharomyces*

cerevisiae during Fermentative and Nonfermentative Growth. *Arch Biochem Biophys* **311**, 62-71
45. Stucka, R., Dequin, S., Salmon, J. M., and Gancedo, C. (1991) DNA sequences in chromosomes II and VII code for pyruvate carboxylase isoenzymes in *Saccharomyces cerevisiae*: analysis of pyruvate carboxylase-deficient strains. *Mol Gen Genet* **229**, 307-315
46. Zhao, J., Zhu, L., Fan, C., Wu, Y., and Xiang, S. (2018) Structure and function of urea amidolyase. *Biosci Rep* **38**
47. Navarathna, D. H. M. L. P., Lionakis, M. S., Lizak, M. J., Munasinghe, J., Nickerson, K. W., and Roberts, D. D. (2012) Urea Amidolyase (*DUR1,2*) Contributes to Virulence and Kidney Pathogenesis of *Candida albicans*. *PLoS One* **7**
48. Kim, H. S., Hoja, U., Stolz, J., Sauer, G., and Schweizer, E. (2004) Identification of the tRNA-binding Protein Arc1p as a Novel Target of *in Vivo* Biotinylation in *Saccharomyces cerevisiae*. *J Biol Chem* **279**, 42445-42452
49. Chang, C.-Y., Chang, C.-P., Chakraborty, S., Wang, S.-W., Tseng, Y.-K., and Wang, C.-C. (2016) Modulating the Structure and Function of an Aminoacyl-tRNA Synthetase Cofactor by Biotinylation *J Biol Chem* **291**, 17102-17111
50. Azhar, A., Booker, G. W., and Polyak, S. W. (2015) Mechanisms of Biotin Transport. *Biochem Anal Biochem* **4**, 1000210-1000211 - 1000210-1000218
51. Stolz, J., Hoja, U., Meier, S., Sauer, N., and Schweizer, E. (1999) Identification of the Plasma Membrane H+-Biotin Symporter of *Saccharomyces cerevisiae* by Rescue of a Fatty Acid-auxotrophic Mutant. *J Biol Chem* **274**, 18741-18746
52. Rogers, T. O., and Lichstein, H. C. (1969) Characterization of the Biotin Transport System in *Saccharomyces cerevisiae*. *J Bacteriol* **100**, 557-564
53. Ahmad Hussin, N. R. A. (2016) *The Roles of Biotin in Candida albicans Physiology.* Degree of Master of Science, University of Nebraska - Lincoln
54. Rogers, T. O., and Lichstein, H. C. (1969) Regulation of Biotin Transport in *Saccharomyces cerevisiae*. *J Bacteriol* **100**, 565-572
55. Becker, J. M., and Lichstein, H. C. (1972) Transport overshoot during biotin uptake by *Saccharomyces cerevisiae Biochim Biophys Acta* **282**, 409-420
56. Phalip, V., Kuhn, I., Lemoine, Y., and Jeltsch, J.-M. (1999) Characterization of the biotin biosynthesis pathway in *Saccharomyces cerevisiae* and evidence for a cluster containing *BIO5,* a novel gene involved in vitamer uptake *Gene* **232**, 43-51
57. Zhang, S., Sanyal, I., Bulboaca, G. H., Rich, A., and Flint, D. H. (1994) The Gene for Biotin Synthase from *Saccharomyces cerevisiase*: Cloning, Sequencing, and Complementation of *Escherichia coli* Strains Lacking Biotin Synthase. *Arch Biochem Biophys* **309**, 29-35
58. Hall, C., and Dietrich, F. S. (2007) The Reacquisition of Biotin Prototrophy in *Saccharomyces cerevisiae* Involved Horizontal Gene Transfer, Gene Duplication and Gene Clustering. *Genetics* **177**, 2293-2307
59. Salaemae, W., Azhar, A., Booker, G. W., and Polyak, S. W. (2011) Biotin biosynthesis in *Mycobacterium tuberculosis*: physiology, biochemistry and molecular intervention. *Protein Cell* **2**, 691-695
60. Firestone, B. Y., and Koser, S. A. (1960) Growth promoting effect of some biotin analogues for *Candida albicans*. *J Bacteriol* **79**, 674-676
61. Ohsugi, M., and Imanishi, Y. (1985) Microbiological Activity of Biotin-Vitamers. *J Nutr Sci Vitaminol* **31**, 563-572
62. Park, S. W., Klotzsche, M., Wilson, D. J., Boshoff, H. I., Eoh, H., Manjunatha, U., Blumenthal, A., Rhee, K., Barry 3rd, C. E., Aldrich, C. C., Ehrt, S., and Schnappinger, D. (2011) Evaluating the Sensitivity of *Mycobacterium tuberculosis* to Biotin Deprivation Using Regulated Gene Expression. *PLoS Pathog* **7**

63. Liu, F., Dawadi, S., Maize, K. M., Dai, R., Park, S. W., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2017) Structure-Based Optimization of Pyridoxal 5'-Phosphate-Dependent Transaminase Enzyme (BioA) Inhibitors that Target Biotin Biosynthesis in *Mycobacterium tuberculosis*. *J Med Chem* **60**, 5507-5520
64. Billones, J. B., Carrillo, M. C. O., Organo, V. G., Sy, J. B. A., Clavio, N. A. B., Macalino, S. J. Y., Emnacen, I. A., Lee, A. P., Ko, P. K. L., and Concepcion, G. P. (2017) In silico discovery and in vitro activity of inhibitors against *Mycobacterium tuberculosis* 7,8-diaminopelargonic acid synthase (*Mtb* BioA). *Drug Des Devel Ther* **11**, 563-574
65. Dai, R., Geders, T. W., Liu, F., Park, S. W., Schnappinger, D., Aldrich, C. C., and Finzel, B. C. (2015) Fragment-Based Exploration of Binding Site Flexibility in *Mycobacterium tuberculosis* BioA. *J Med Chem* **58**, 5208-5217
66. Park, S. W., Casalena, D. E., Wilson, D. J., Dai, R., Nag, P. P., Liu, F., Boyce, J. P., Bittker, J. A., Schreiber, S. L., Finzel, B. C., Schnappinger, D., and Aldrich, C. C. (2015) Target-Based Identification of Whole-Cell Active Inhibitors of Biotin Biosynthesis in *Mycobacterium tuberculosis*. *Chem Biol* **22**, 76-86
67. Shi, C., and Aldrich, C. C. (2012) Design and Synthesis of Potential Mechanism-Based Inhibitors of the Aminotransferase BioA Involved in Biotin Biosynthesis. *J Org Chem* **77**, 6051-6058
68. Shi, C., Geders, T. W., Park, S. W., Wilson, D. J., Boshoff, H. I., Abayomi, O., Barry 3rd, C. E., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2011) Mechanism-based Inactivation by Aromatization of the Transaminase BioA Involved in Biotin Biosynthesis in *Mycobacterium tuberculosis*. *J Am Chem Soc* **133**, 18194-18201
69. Mann, S., Colliandre, L., Labesse, G., and Ploux, O. (2009) Inhibition of 7,8-diaminopelargonic acid aminotransferase from *Mycobacterium tuberculosis* by chiral and achiral anologs of its substrate: Biological implications. *Biochimie* **91**, 826-834
70. Zakikhany, K., Naglik, J. R., Schmidt-Westhausen, A., Holland, G., Schaller, M., and Hube, B. (2007) *In vivo* transcript profiling of *Candida albicans* identifies a gene essential for interepithelial dissemination. *Cell Microbiol* **9**, 2938-2954
71. Hoja, U., Wellein, C., Greiner, E., and Schweizer, E. (1998) Pleiotropic phenotype of acetyl-CoA-carboxylase-defective yeast cells: Viability of a BPL1-amber mutation depending on its readthrough by normal tRNA (Gln)(CAG). *Eur J Biochem* **254**, 520-526
72. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Rhonda, B., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., Bakkoury, M. E., Foury, F., Friend, S. H., Gentalen, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D. J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J. L., Riles, L., Roberts, C. J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R. K., Véronneau, S., Voet, M., Volckaert, G., Ward, T. R., Wysocki, R., Yen, G. S., Yu, K., Zimmermann, K., and Philippsen, P. (1999) Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science* **285**, 901-906
73. Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H.-S., Duhig, T., Nam, M., Palmer, G., Han, S., Jeffery, L., Baek, S.-T., Lee, H., Shim, Y. S., Lee, M., kIM, L., Heo, K.-S., Noh, E. J., Lee, A.-R., Jang, Y.-J., Chung, K.-S., Choi, S.-J., Park, J.-Y., Park, Y., Kim, H. M., Park, S.-K., Park, H.-J., Kang, E.-J., Kim, H. B., Kang, H.-S., Park, H.-M., Kim, K., Song, K., Song, K. B., Nurse, P., and Hoe, K.-L. (2010) Analysis of a genome-wide set of gene deletions in the fissuion yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* **28**, 617-623
74. Shtifman Segal, E., Gritsenko, V., Levitan, A., Yadav, B., Dror, N., Steenwyk, J. L., Silberberg, Y., Mielich, K., Rokas, A., Gow, N. A. R., Kunze, R., Sharan, R., and

Berman, J. (2018) Gene Essentiality Analyzed by *In Vivo* Transposon Mutagenesis and Machine Learning in a Stable Haploid Isolate of *Candida albicans*. *mBio* **9**
75. Yao, T.-T., Xiao, D.-X., Li, Z.-S., Cheng, J.-L., Fang, S.-W., Du, Y.-J., Zhao, J.-H., Dong, X.-W., and Zhu, G.-N. (2017) Design, Synthesis, and Fungicidal Evaluation of Novel Pyrazole-furan and Pyrazole-pyrrole Carboxamide as Succinate Dehydrogenase Inhibitors. *J Agric Food Chem* **65**, 5397-5403
76. Fraaije, B. A., Bayon, C., Atkins, S., Cools, H. J., Lucas, J. A., and Fraaije, M. W. (2012) Risk assessment studies on succinate dehydrogenase inhibitors, the new weapons in the battle to control Septoria leaf blotch in wheat. *Mol Plant Pathol* **13**, 263-275
77. Bockman, M. R., Engelhart, C. A., Dawadi, S., Larson, P., Tiwari, D., Ferguson, D. M., Schnappinger, D., and Aldrich, C. C. (2018) Avoiding Antibiotic Inactivation in *Mycobacterium tuberculosis* by Rv3406 through Strategic Nucleoside Modification. *ACS Infect Dis* **4**, 1102-1113
78. Hayes, A. J. (2017) *Biotin protein ligase inhibitors as new antibacterial agents to target Staphylococcus aureus: Studies of efficacy, mechanism of action and resistance.* Master of Philosophy, University of Adelaide
79. Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40
80. Gerdes, S. Y., Scholle, M. D., Campbell, J. W., Balázsi, g., Ravasz, E., Daugherty, M. D., Somera, A. L., Kyrpides, N. C., Anderson, I., Gelfand, M. S., Bhattacharya, A., Kapatral, V., D'Souza, M., Baev, M. V., Grechkin, Y., Mseeh, F., Fonstein, M. Y., Overbeek, R., Barabási, A.-L., Oltvai, Z. N., and Osterman, A. L. (2003) Experimental Determination and System Level Analysis of Essential Genes in *Escherichia coli* MG1655. *J Bacteriol* **185**, 5673-5684
81. Duckworth, B. P., Geders, T. W., Tiwari, D., Boshoff, H. I., Sibbald, P. A., Barry III, C. E., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2011) Bisubstrate Adenylation Inhibitors of Biotin Protein Ligase from *Mycobacterium tuberculosis*. *Chem Biol* **18**, 1432-1441
82. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Zvarec, O., Bell, J. M., Turnidge, J. D., Wallace, J. C., Booker, G. W., Wilce, M. C. J., Abell, A. D., and Polyak, S. W. (2012) Biotin Analogues with Antibacterial Activity Are Potent Inhibitors of Biotin Protein Ligase. *ACS Med Chem Lett* **3**, 509-514
83. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Pendini, N. R., Polyak, S. W., Pedersen, D. S., Morona, R., Turnidge, J. D., Wallace, J. C., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2012) Selective inhibition of Biotin Protein Ligase from *Staphylococcus aureus*. *J Biol Chem* **287**, 17823-17832
84. Paparella, A. S., Lee, K. J., Hayes, A. J., Feng, J., Feng, Z., Cini, D., Deshmukh, S., Booker, G. W., Wilce, M. C. J., Polyak, S. W., and Abell, A. D. (2018) Halogenation of Biotin Protein Ligase Inhibitors Improves Whole Cell Activity against *Staphylococcus aureus*. *ACS Infect Dis* **4**, 175-184
85. Feng, J., Paparella, A. S., Tieu, W., Heim, D., Clark, S., Hayes, A., Booker, G. W., Polyak, S. W., and Abell, A. D. (2016) New Series of BPL Inhibitors To Probe the Ribose-Binding Pocket of *Staphylococcus aureus* Biotin Proteing Ligase *ACS Med Chem Lett* **7**, 1068-1072
86. Tieu, W., Polyak, S. W., Paparella, A. S., Yap, M. Y., Soares da Costa, T. P., Ng, B., Wang, G., Lumb, R., Bell, J. M., Turnidge, J. D., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2015) Improved Synthesis of Biotinol-5′-AMP: Implications for Antibacterial Discovery. *ACS Med Chem Lett* **6**, 216-220
87. Tieu, W., Jarrad, A. M., Paparella, A. S., Keeling, K. A., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., Polyak, S. W., and Abell, A. D. (2014) Heterocyclic

acyl-phosphate bioisostere-based inhibitors of *Staphylococcus aureus* biotin protein ligase. *Bioorg Med Chem Lett* **24**, 4689-4693

88. Bockman, M. R., Kalinda, A. S., Petrelli, R., De la Mora-Rey, T., Tiwari, D., Liu, F., Dawadi, S., Nandakumar, M., Rhee, K. Y., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2015) Targeting *Mycobacterium tuberculosis* Biotin Protein Ligase (MtBPL) with Nucleoside-Based Bisubstrate Adenylation Inhibitors. *J Med Chem* **58**, 7349-7369
89. Tiwari, D., Park, S. W., Essawy, M. M., Dawadi, S., Mason, A., Nandakumar, M., Zimmerman, M., Mina, M., Pin Ho, H., Engelhart, C. A., Ioerger, T., Sacchettini, J. C., Rhee, K., Ehrt, S., Aldrich, C. C., Dartois, V., and Schnappinger, D. (2018) Targeting protein biotinylation enhances tuberculosis chemotherapy. *Sci Transl Med* **10**
90. Mayende, L. (2012) *Functional characterisation of the N-terminal region of Holocarboxylase synthetase.* Doctor of Philosophy, University of Adelaide

CHAPTER 3:

DNA-BINDING PATHWAYS OF *Staphylococcus aureus* BPL

CHAPTER OUTLINE

This chapter contains a published manuscript that I contributed to as an equal first author. Here I optimized a native mass spectrometry (MS) assay to study a previously well-characterised BPL, namely the *Staphylococcus aureus* BPL (*Sa*BPL). This approach was then utilised, in tandem with orthogonal techniques, to study the self-oligomerisation and DNA-binding activity of a mutant *Sa*BPL in comparison to the wild-type enzyme, and these data revealed novel insights into the DNA-binding pathways of *Sa*BPL. This work established the native MS technique that was later used in chapters 4 and 5 for the study of fungal class III BPLs.

STATEMENT OF AUTHORSHIP

Statement of Authorship

Title of Paper	Native mass spectrometry identifies an alternative DNA-binding pathway for BirA from *Staphylococcus aureus*
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Published in Scientific Reports, Feb 2019, Vol 9, Doi: 10.1038/s41598-019-39398-6

Principal Author

Name of Principal Author (Candidate)	Louise Sternicki
Contribution to the Paper	Prepared and analysed samples by native mass spectrometry, analysed mass spectrometry data, prepared manuscript and figures
Overall percentage (%)	40%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 6/3/2019.

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Jiulia Satiaputra
Contribution to the Paper	Recombinant production of proteins, EMSA DNA binding assays, developed LacZ reporter constructs and completed assays, data analysis, prepared manuscript
Signature	Date 15/05/2019

Name of Co-Author	Andrew Hayes
Contribution to the Paper	Sequence aiignments of different BirA sequences across *Staphylococcus* strains
Signature	Date 22/04/19

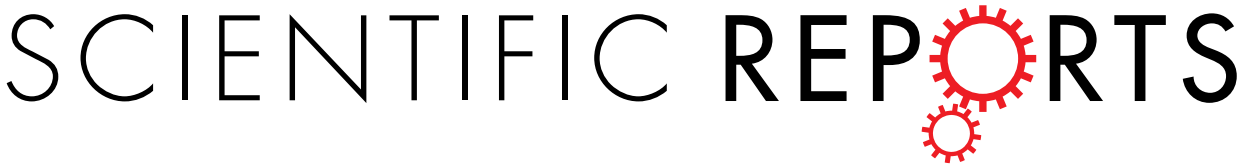
Name of Co-Author	Tara Pukala		
Contribution to the Paper	Supervised mass spectrometry data collection and analysis, helped with manuscript preparation		
Signature		Date	6/3/19

Name of Co-Author	Grant Booker		
Contribution to the Paper	Intellectual discussion of the project, prepare manuscript		
Signature		Date	6/3/2019

Name of Co-Author	Keith Shearwin		
Contribution to the Paper	Intellectual discussion of the project, prepare manuscript		
Signature		Date	6/3/2019

Name of Co-Author	Steven Polyak		
Contribution to the Paper	Conceived the study, intellectual discussion of the project, prepare manuscript, corresponding author		
Signature		Date	6/3/2019

www.nature.com/scientificreports



OPEN

Native mass spectrometry identifies an alternative DNA-binding pathway for BirA from *Staphylococcus aureus*

Received: 14 September 2018
Accepted: 15 January 2019
Published online: 26 February 2019

Jiulia Satiaputra[1,3], Louise M. Sternicki[1], Andrew J. Hayes[1,4], Tara L. Pukala[2], Grant W. Booker[1], Keith E. Shearwin[1] & Steven W. Polyak[1,5]

An adequate supply of biotin is vital for the survival and pathogenesis of *Staphylococcus aureus*. The key protein responsible for maintaining biotin homeostasis in bacteria is the biotin retention protein A (BirA, also known as biotin protein ligase). BirA is a bi-functional protein that serves both as a ligase to catalyse the biotinylation of important metabolic enzymes, as well as a transcriptional repressor that regulates biotin biosynthesis, biotin transport and fatty acid elongation. The mechanism of BirA regulated transcription has been extensively characterized in *Escherichia coli*, but less so in other bacteria. Biotin-induced homodimerization of *E. coli* BirA (*Ec*BirA) is a necessary prerequisite for stable DNA binding and transcriptional repression. Here, we employ a combination of native mass spectrometry, *in vivo* gene expression assays, site-directed mutagenesis and electrophoretic mobility shift assays to elucidate the DNA binding pathway for *S. aureus* BirA (*Sa*BirA). We identify a mechanism that differs from that of *Ec*BirA, wherein *Sa*BirA is competent to bind DNA as a monomer both in the presence and absence of biotin and/or MgATP, allowing homodimerization on the DNA. Bioinformatic analysis demonstrated the *Sa*BirA sequence used here is highly conserved amongst other *S. aureus* strains, implying this DNA-binding mechanism is widely employed.

The biotin retention protein A (BirA, also known as biotin protein ligase) is responsible for maintaining biotin homeostasis in many bacteria. BirA is a bi-functional protein capable of both enzymatic protein biotinylation as well as serving as a biotin-controlled transcriptional repressor that regulates the expression of the biotin biosynthesis operon (reviewed[1,2]). By combining both activities in a single protein, BirA is uniquely placed as the key regulator of biotin metabolism. *E. coli* BirA (*Ec*BirA) has been thoroughly investigated through genetic, biochemical and structural biology studies[3–11]. These have provided powerful insights into the maintenance of biotin homeostasis. *Ec*BirA binds to its ligands, biotin and ATP, in an ordered manner[6,12]. Conformational changes induced by biotin binding create the pocket necessary for ATP to bind. ATP binding initiates the synthesis of biotinyl-5′-AMP, which serves as both a reaction intermediate for biotin ligation as well as a co-repressor. The BirA:biotinyl-5′-AMP complex, known as the holo-enzyme, is then available for two alternative activities: either as a biotin protein ligase when there is a substrate available for protein biotinylation or when demand for biotin is low, as a transcriptional repressor through homodimerization[13]. The transcriptional repressor function of *Ec*BirA involves a co-operative interaction between two *Ec*BirA subunits and an inverted palindromic repeat sequence present in the promoter of the biotin biosynthetic operon (*bioO*). Homodimerization of the holo-enzyme complex ($K_D^{2\text{-}1} = 1\text{–}6 \times 10^{-6}$ M[14,15]) is a pre-requisite for DNA binding[4,5,16,17]. The unliganded enzyme (i.e. apo-*Ec*BirA) does not dimerize at physiological concentrations ($K_D^{2\text{-}1} = 1 \times 10^{-3}$ M[14]) and is unable to bind DNA[9,17]. Mutagenesis studies have further highlighted the importance of homodimerisation in the DNA binding

[1]School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005, Australia. [2]School of Physical Sciences, University of Adelaide, Adelaide, South Australia, 5005, Australia. [3]Present address: Harry Perkins Institute of Medical Research, Shenton Park, Western Australia, 6008, Australia. [4]Present address: Faculty of Health and Medical Sciences, Adelaide, South Australia, 5005, Australia. [5]Present address: School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, South Australia, 5001, Australia. Jiulia Satiaputra and Louise M. Sternicki contributed equally. Correspondence and requests for materials should be addressed to S.W.P. (email: steven.polyak@adelaide.edu.au)

mechanism. Amino acid substitutions that promote homo-dimerization display increased affinity for DNA and function as super-repressors[18,19]. Conversely, replacing key amino acids in the interface between the two *Ec*BirA subunits disrupts dimerization, and eliminates DNA-binding activity[7,15,17,20,21]. One well-characterized example is *Ec*BirA-R119W[20,22] (Supporting Fig. S1A) that is essentially monomeric in solution ($K_D^{2\text{-}1} = 2 \times 10^{-2}$ M[17]). This dimerization-impaired mutant thus mimics the association state of apo-*Ec*BirA.

S. aureus BirA (*Sa*BirA) shares many features with its *E. coli* homologue. The two proteins have similar tertiary structures despite limited sequence conservation (24% identity, 42% similarity, Supporting Fig. S1B). Both proteins undergo analogous conformational changes that define the ordered ligand binding mechanism[23–25]. Like *Ec*BirA, *Sa*BirA is a biotin-dependent transcriptional repressor of biotin biosynthesis[26,27]. Unlike *Ec*BirA however, *Sa*BirA regulates expression of additional operons - the substrate specific subunit of the biotin transporter (encoded by *bioY*), and the *yhfS-yhfT* operon encoding putative homologs of acetyl-CoA acetyl transferase and long-chain fatty acid-CoA ligase, respectively[26,28]. Holo-*Sa*BirA binding directly to the promoter sequence of the biotin biosynthesis operon (*SabioO*) has recently been demonstrated by us[24,26,29], and others[27]. Estimates of the dissociation constant (K_D), based on polyacrylamide gel-based electrophoretic mobility shift assays (EMSA), suggest a 6-fold difference in the affinity for DNA between apo and holo-*Sa*BirA (apo 649 nM, holo 108 nM[29]). In support of these results, fluorescence anisotropy experiments performed in solution suggest the holo-protein binds 63-fold tighter to DNA than the apo-protein (apo 5 μM, holo 83 nM[27]). The modest differences in affinity for the apo enzyme are likely due to polymorphisms in the *Sa*BirA sequences used in the two studies. Despite the intrinsic DNA-binding activity of the apo enzyme, *Sa*BirA has been shown to function as a biotin-dependent transcriptional repressor[26,27]. These observations are significantly different to *Ec*BirA, where apo-protein is devoid of DNA-binding activity.

In this study, the homodimerisation and DNA-binding activities of *Ec*BirA and *Sa*BirA were investigated using *in vivo* reporter assays and *in vitro* biochemical analysis. For the first time, native mass spectrometry was utilized to study the self-association and DNA binding functions of BirA. Nano-electrospray ionization-mass spectrometry (nESI-MS) performed under native conditions is an ideal technique for studying BirA as the soft ionization employed maintains non-covalent interactions during the transit from solution to gas phase[30]. This permitted the direct detection of biotinyl-5′-AMP binding to BirA, as well as quantitating the stoichiometry of the protein:DNA complexes. Engineered protein mutants designed to disrupt homodimerization, namely *Ec*BirA-R119W and the *S. aureus* equivalent *Sa*BirA-F123G (Supporting Fig. S1), were also employed to address the requirement for protein dimerization on DNA-binding (*in vitro*) and transcriptional repression (*in vivo*). Together these data revealed that monomeric *Sa*BirA is competent to bind DNA and, once bound, can recruit a second protein subunit. This new pathway is not evident with *Ec*BirA. Bioinformatic analysis identified that the *Sa*BirA enzyme used in this study, from the methicillin and vancomycin resistant *S. aureus* strain Mu50, is the prototypical BirA amongst the *S. aureus* species. We propose that this DNA binding mechanism may be widely used amongst *S. aureus*.

Results

Biotin-induced repression of transcription. Apo-*Sa*BirA has been shown to bind DNA *in vitro*[27,29]. We sought to address if the non-liganded protein possessed repressor activity in an *in vivo* assay of gene expression. As it is technically challenging to achieve the apo-state *in vivo*, we employed the *Sa*BirA-F123G mutant that impairs homo-dimerization and, therefore, mimics the monomeric properties of the non-liganded protein[29]. Wild-type *Ec*BirA and *Sa*BirA, and the dimerization disrupted *Ec*BirA-R119W mutant[31], were used as controls. Biotin-induced repressor activity was measured using a bacterial reporter system we have previously described[26] and summarized in Fig. 1A. Briefly, a *birA* gene under the control of the pLac-UV5 promoter was site-specifically integrated into the lambda phage attachment site (attB) in the chromosome of a biotin auxotrophic strain of *E. coli* harboring a mutant BirA that is unable to bind DNA[26]. The host strain also produces lac repressor from the strong $LacI^q$ promoter, allowing us to repress expression of the BirA variants down to the low concentrations necessary for a physiologically relevant assay. A second construct containing a BirA-target promoter fused to a *lacZ* reporter gene was then integrated into the HK022 attB phage attachment site. Here, the promoters for *bioO* from both *E. coli* (*EcbioO*) and *S. aureus* (*SabioO*) were analyzed alongside promoters for the *bioY* and the *yhfS-yhfT* operon in *S. aureus*[26]. Biotin-mediated gene control was then assayed using β-galactosidase activity as a readout for gene expression. The effect of exogenous biotin on these genetic circuits, engineered into the genome of a strain unable to synthesis its own biotin, was measured by the addition of biotin to the growth media.

The *in vivo* reporter assay was developed to reconstitute the biotin-induced transcriptional repressor activity of BirA. Previous studies have suggested that *E. coli* contains ≈10–20 molecules of BirA per cell[32]. Therefore, the assay was designed to minimize artefactual binding to DNA due to the overexpression of BirA. Our previous work has demonstrated that the level of lac repressor provided from a single copy $pLacI^q$-lacI module is able to very significantly, though not completely, repress a single copy plac-UV5 promoter module[33]. As expected, biotin-regulated repression of β-galactosidase was observed when IPTG was omitted from the growth media - implying the low level of leaky expression from the repressed lacUV5 promoter provided sufficient *Sa*BirA for activity. Conversely, addition of 10 μM IPTG induced much higher expression of *Sa*BirA, as detected by Western blot, and also abolished biotin-regulated control of the reporter (Supporting Fig. S2). Similarly, the level of *Ec*BirA expression obtained in the absence of IPTG was shown to progressively inhibit β-galactosidase expression with increasing levels of biotin present in the growth media, as expected (Fig. 1B). Hence, all *in vivo* experiments were performed in the absence of IPTG (Fig. 1B–E).

The concentrations of biotin required to achieve half-maximum repression at equilibrium (R_{bio}) were calculated for each protein (Table 1). For *E. coli*, half maximal repression was achieved with 3.8 nM biotin. Likewise, biotin-induced repression was similar for all three *S. aureus* target promoters, with 7–10 nM biotin required for the same level of response (Table 1). Consistent with previous literature, the *Ec*BirA-R119W mutant that lacks homo-dimerisation activity[22,31] was also devoid of repressor activity *in vivo* (Fig. 1B, green symbols) yielding

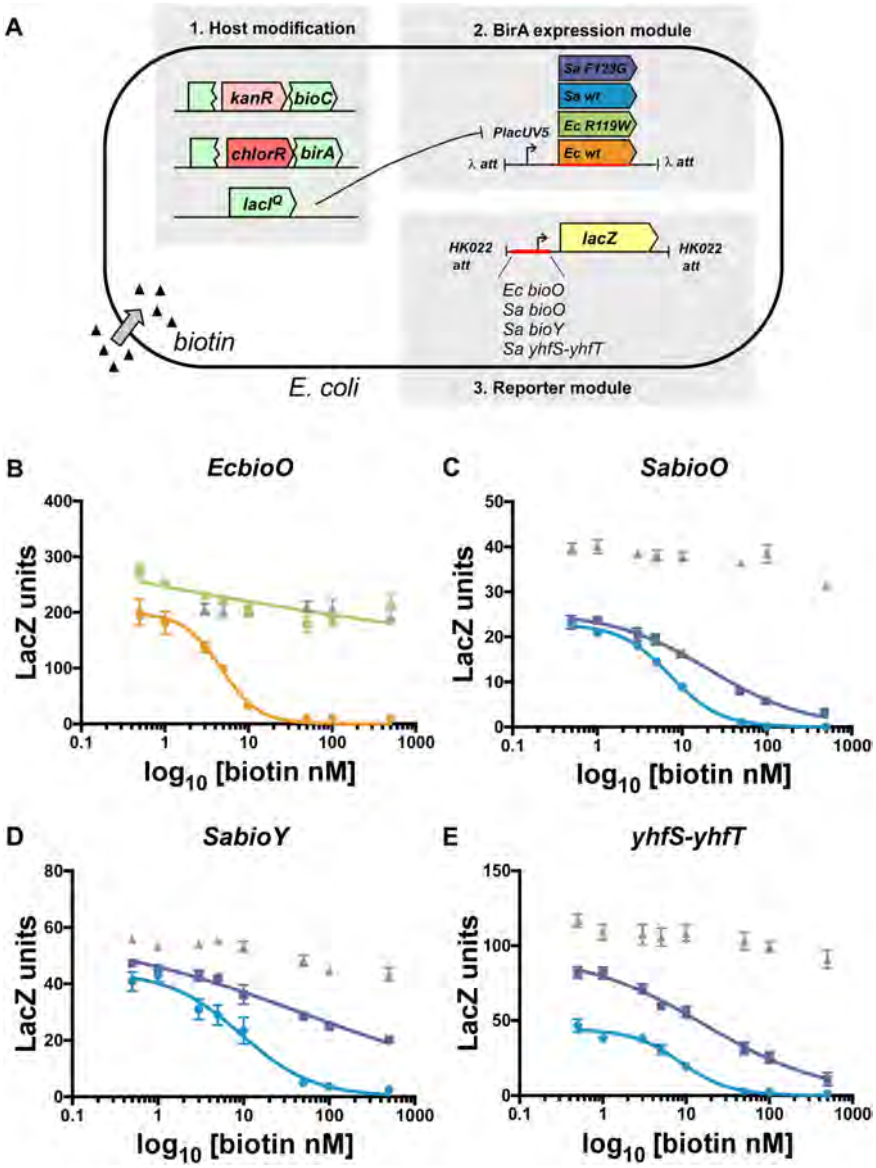


Figure 1. *In vivo* β-galactosidase assays to measure biotin-induced repression. (**A**) Overview of *E. coli* reporter strains containing chromosomally integrated reporters. (**B**–**E**) β-galactosidase assays revealing biotin-induced repression by *Ec*BirA (orange circles), *Ec*BirA-R119W (green squares), *Sa*BirA (blue circles) or *Sa*BirA-F123G (purple squares). Target promoter sequences investigated were (**B**) *EcbioO*, (**C**) *SabioO*, (**D**) *SabioY* or (**E**) *SayhfS-yhfT*. Strains with no integrated BirA served as controls (grey triangles). A further control lacking an integrated promoter was used to measure the background *lacZ* activity at each biotin concentration (≤10 units), which has been subtracted to give values shown in the graphs. Error bars denote S.E.M from independent biological replicates (n = 6).

similar levels of β-galactosidase expression as the no-repressor control (Fig. 1B, grey symbols). In contrast, *Sa*BirA-F123G repressed activity against all three target promoters in a biotin-dependent manner (Fig. 1C–E, purple curves). For *SabioO* and *SayhfS-yhfT*, the R_{bio} values were 2.6-fold and 2.3-fold higher for *Sa*BirA-F123G than wild-type protein, respectively (Table 1) ($p = 0.07$ WT *vs* F123G for *SabioO*; $p = 0.2$, WT *vs* F123G for *SayhfS-yhfT*). This modest decrease in repressor activity is consistent with a 3-fold higher K_M for biotin that has been reported for *Sa*BirA-F123G relative to the wild-type enzyme[29]. For *SabioY*, an accurate estimate of R_{bio} was not achieved as 500 nM of biotin was insufficient to completely inhibit expression down to the background level necessary to generate a concentration-dependent repression curve. At the lowest biotin concentration tested (1 nM), β-galactosidase expression from the *SabioO* and *SayhfS-yhfT* promoters was significantly lower than the

Promoter	Repression constant (R_{bio}) nM	
	Wild-type	Dimerization-impaired mutant
EcbioO	3.8 ± 0.6	>500
SabioO	7.0 ± 0.3	18.6 ± 5.6
SabioY	6.1 ± 2.6	>500
SayhfS-yhfT	10.0 ± 2.0	22.6 ± 9.7

Table 1. *In vivo* equilibrium binding constants for biotin-induced transcriptional repression. The concentration of biotin required to achieve half-maximum repression (R_{bio}) at equilibrium in the β-galactosidase reporter assay was calculated from the data presented in Fig. 1B–E. Data was fitted using the one-site specific binding equation $Y = Bmax*X^h/(R_{bio}{}^h + X^h)$, where Y = *LacZ* units, X = biotin concentration (nM), Bmax = maximum binding (*LacZ* units) and h = Hill-slope (BirA binding as a homodimer, h = 2). Datum is from six independent experiments.

corresponding no-repressor controls (Fig. 1C,E, wild-type blue and *Sa*BirA-F123G purple *vs* control grey symbols) implying both the wild-type and *Sa*BirA-F123 mutant proteins partially occupied the DNA under these conditions. Together this data suggests monomeric *Sa*BirA (a mimic of the apo-state) is capable of repressing its target genes, which is not evident with *Ec*BirA.

***In vitro* analysis of apo and holo-BirA oligomeric states.** The molecular basis of the unexpected repressor activity observed with *Sa*BirA-F123G was further investigated using biochemical assays. Wild-type *Ec*BirA and *Sa*BirA were purified in their apo-forms alongside the two dimerization-impaired mutant proteins *Ec*BirA-R119W and *Sa*BirA-F123G. Removal of biotin from the recombinant proteins was achieved by incubating cell lysates (containing over-expressed protein) with streptavidin-Sepharose to remove excess biotin, before incubation with a biotin-accepting substrate protein to facilitate protein biotinylation and the concomitant loss of biotinyl-5′-AMP from the BirA active site prior to IMAC purification. The apo-state of the four proteins was addressed using two alternative techniques that independently confirmed that all preparations were devoid of any co-purified biotinyl-5′-AMP; a streptavidin-blot method we have previously described[29] (Supporting Fig. S3) and native nESI-MS that has not been employed previously for the study of BirA. The Streptavidin-blot assay failed to detect biotinyl-transferase activity for all four proteins, as expected for an apo-enzyme. In support, the masses measured by nESI-MS for the apo-preparations were as expected for monomeric BirA devoid of ligand (Fig. 2A,C,E and G and Supporting Table S4). We confirmed that the mild conditions employed with the native mass spectrometry studies allowed the direct measurement of BirA in complex with biotinyl-5′-AMP, as well as the oligomeric state of holo-BirA (Fig. 2B,D,F and H). The addition of biotin and MgATP to *Ec*BirA and *Sa*BirA resulted in the detection of two species (Fig. 2B,F), consistent with those expected for the monomeric proteins bound to the reaction intermediate biotinyl-5′-AMP (*Ec*BirA, measured 36771 Da, theoretical molecular mass 36765 Da: *Sa*BirA, measured 38470 Da, theoretical molecular mass 38466 Da) and the dimeric form of these ligand-bound complexes (*Ec*BirA, measured 73559 Da, theoretical molecular mass 73530 Da: *Sa*BirA: measured 76925 Da, theoretical molecular mass 76931 Da). The nESI-MS data confirmed that *Ec*BirA and *Sa*BirA possessed biotinyl-5′-AMP synthetase activity, and the enzyme:biotinyl-5′-AMP complex was stable following their exchange into the volatile 200 mM ammonium acetate pH 6.85 buffer required for mass spectrometry. Incubation of *Ec*BirA-R119W and *Sa*BirA-F123G with biotin and MgATP yielded single species with masses consistent with monomers in complex with biotinyl-5′-AMP (*Ec*BirA-R119W, measured 36783 Da, theoretical molecular mass 36795 Da: *Sa*BirA-F123G, measured 38381 Da, theoretical molecular mass of 38376 Da). There was no evidence of biotin-induced homodimers for either mutant protein (Fig. 2D,H). This demonstrated that whilst *Ec*BirA-R119W and *Sa*BirA-F123G retained biotinyl-5′-AMP synthetase activity, both had abolished in-solution dimerization activity.

Apo and dimerization-impaired *Sa*BirA are competent to bind DNA. Native nESI-MS (Fig. 3) was employed to probe the DNA binding activities of *Ec*BirA and *Sa*BirA. Oligonucleotides containing the recognition sequences from the *EcbioO* and *S. aureus bioY* promoters were employed for the nESI-MS analysis. DNA:protein complexes were formed in 200 mM ammonium acetate pH 6.9 buffer by incubating BirA with the double-stranded oligonucleotides on ice for at least one hour prior to mass spectrometry analysis. A comparison of the theoretical and measured molecular masses for the complexes analysed in this study is shown in Supporting Table S5. The specificity of the BirA:DNA interaction was initially confirmed using an oligonucleotide lacking BirA operators. Neither native nESI-MS nor EMSA detected *Sa*BirA binding to an oligonucleotide with the two BirA half sites mutated to random sequences (Supporting Fig. S4).

For holo-*Ec*BirA, the predominant species had a mass consistent with two subunits of *Ec*BirA bound to DNA. For the apo-*Ec*BirA protein, this species was seen only with weak intensity (Fig. 3A,B). For *Ec*BirA-R119W (at a protein concentration of 10 μM) biotinyl-5′-AMP also induced a complex comprising two protein subunits on DNA, a species which was only observed with a weak intensity for apo-*Ec*BirA-R119W (Fig. 3C,D). At equilibrium, the majority of the apo-wild-type *Ec*BirA and *Ec*BirA-R119W was present as monomeric protein, not in complex with DNA. Consistent with the nESI-MS datum, DNA binding was also observed in an EMSA for wild-type holo *Ec*BirA (Fig. 4A). However, holo-*Ec*BirA-R119W mutant binding to DNA could not be measured by EMSA, even at 400 nM BirA (Fig. 4B).

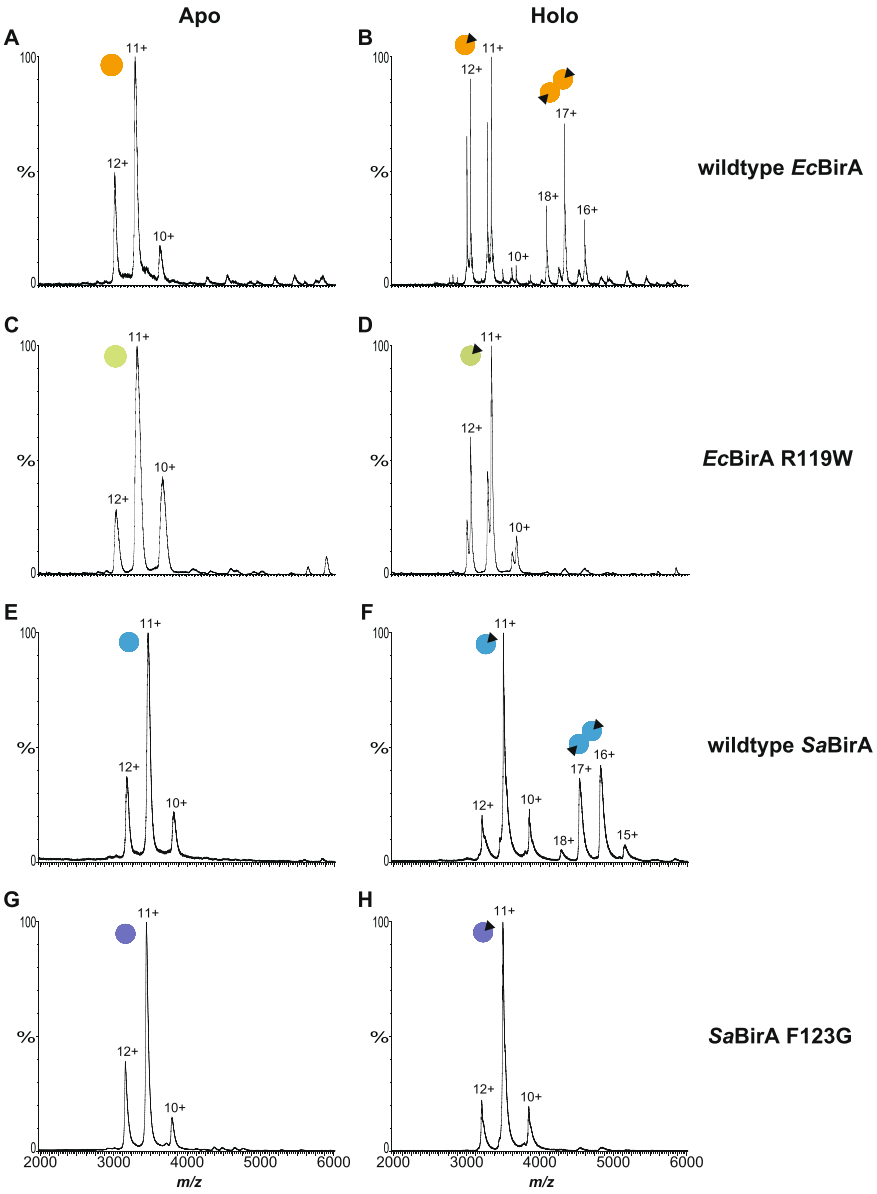
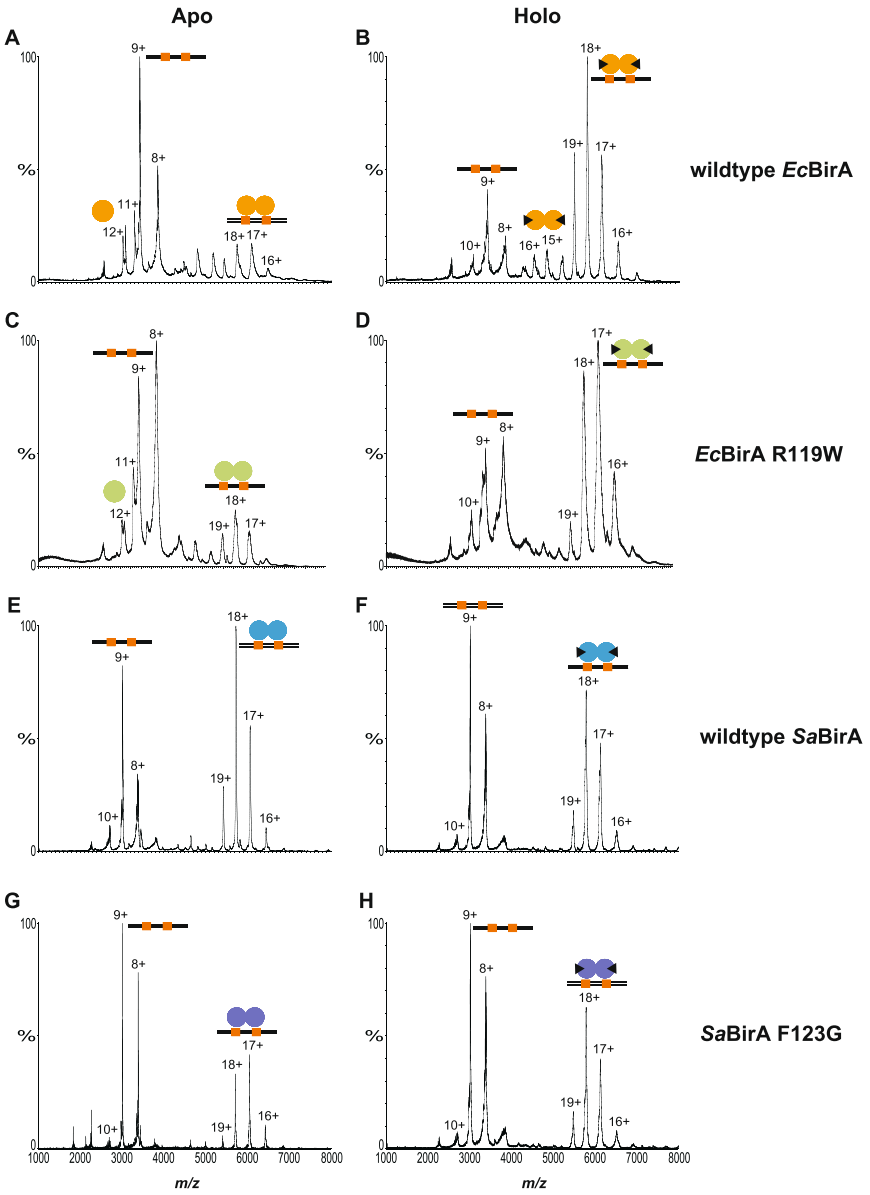


Figure 2. Native nano-electrospray ionization-mass spectrometry determination of the oligomeric state of wild-type BirA and dimerization-impaired mutants. nESI-MS spectra of (**A**) apo-*Ec*BirA, (**B**) holo-*Ec*BirA, (**C**) apo-*Ec*BirA-R119W, (**D**) holo-*Ec*BirA-R119W, (**E**) apo-*Sa*BirA, (**F**) holo-*Sa*BirA, (**G**) apo-*Sa*BirA-F123G and (**H**) holo-*Sa*BirA-F123G. Peaks revealing the oligomeric state of the proteins are marked by the sphere symbols and are annotated with the corresponding charge states. Monomers are denoted by a single sphere, homodimers are shown by two joined-spheres and the presence of biotinyl-5′-AMP is represented with a black triangle.

For the *Sa*BirA, evidence of two protein subunits in complex with DNA was observed for wild-type and mutant *Sa*BirA-F123G, in both apo and holo-states (Fig. 3E–H). In both apo-samples minimal free protein was observed, implying most of the *Sa*BirA was in a complex with DNA. This was in contrast to apo-*Ec*BirA where free protein was readily detected. This datum was consistent with EMSA results where both holo-wild-type and holo-*Sa*BirA-F123G were competent to bind DNA (Fig. 4C–H). The EMSA analysis also revealed the holo-wild-type *Sa*BirA binds all three target promoters at similar affinities (Fig. 4C–E). However, the interaction between holo-*Sa*BirA-F123G and *Sa yhfS-yhFT* was clearly weaker compared to the *SabioO* and *SabioY* operator sequences (Fig. 4E–G) which is likely due to a base pair mismatch in the BirA binding site[26]. Together these findings highlight a key difference between the *E. coli* and *S. aureus* BirAs – monomeric (i.e. apo) *Sa*BirA is capable of stable binding to DNA.



***Ec*bioO:** TAATCGAC**TTGTAAACC**AAATTGAAAAGATTTA**GGTTTACAA**GTCTACAC
ATTAGCTG**AACATTTGG**TTTAACTTTTCTAAAT**CCAAATGTT**CAGATGTG

***Sa*bioY:** AACTT**ATTGTAAAC**TTTTCATTTCTTAAA**GTTTACAAT**GGTGCT
TTGAA**TAACATTTG**AAAAGTAAAGAATTT**CAAATGTTA**CCACGA

Figure 3. Native nano-electrospray ionization-mass spectrometry analysis of BirA binding to DNA. nESI-MS spectra demonstrating the protein:DNA interaction involving (**A**) apo-*Ec*BirA, (**B**) holo-*Ec*BirA, (**C**) apo-*Ec*BirA-R119W and (**D**) holo-*Ec*BirA-R119W bound to *EcbioO* and (**E**) apo-*Sa*BirA, (**F**) holo-*Sa*BirA, (**G**) apo-*Sa*BirA-F123G and (**H**) holo-*Sa*BirA-F123G bound to *SabioY*. Monomers are denoted by a single sphere, homodimers are shown by two joined-spheres and the presence of biotinyl-5′-AMP is represented with a black triangle. Black lines represent double stranded oligonucleotides containing two BirA binding sites (orange). Peaks corresponding to the major species are annotated with their charge state. The DNA sequences of the oligonucleotides used in the assay are shown below with the two half sites highlighted in red.

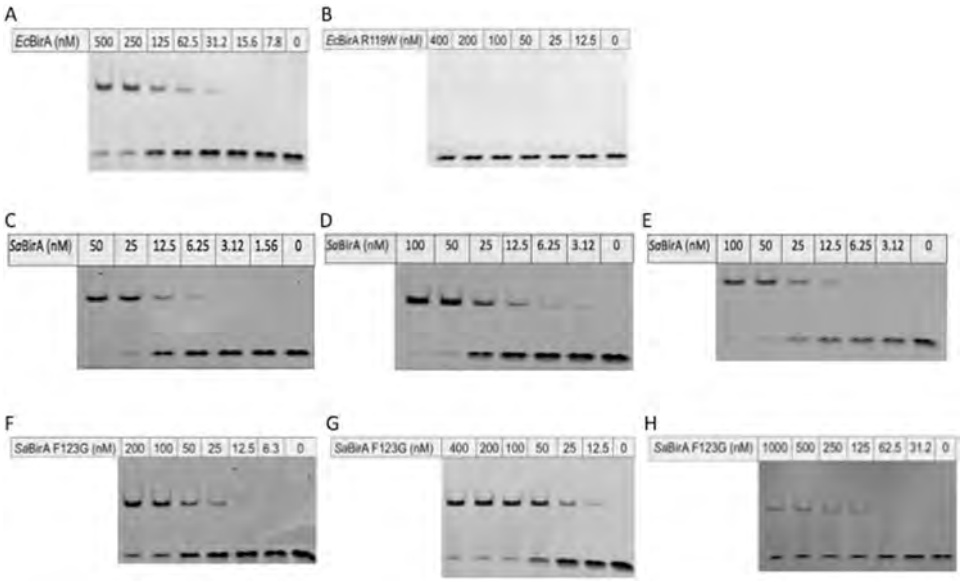


Figure 4. Electrophoretic mobility shift assays. EMSA of the interaction between *EcbioO* and (**A**) *Ec*BirA or (**B**) *Ec*BirA-R119W. Also shown is *Sa*BirA binding to (**C**) *SabioO*, (**D**) *SabioY* or (**E**) *SayhfS-yhfT* and *Sa*BirA-F123G binding to (**F**) *SabioO*, (**G**) *SabioY* and (**E**) *SayhfS-yhfT*. Concentrations of enzyme used in the binding reactions are indicated. Panels C, D and E showing wild type *Sa*BirA binding are from[26]. Full length gels are presented in Supporting Fig. S5.

Ordered assembly of BirA on DNA. The ordered assembly of *Sa*BirA on DNA was finally addressed by native nESI-MS using a *SabioO* oligonucleotide that contained a single half site (Fig. 5, Supporting Table S7). The apo-*Sa*BirA sample revealed a species consistent with one monomer bound to DNA with a charge state distribution centered around 13+ and 14+ (Fig. 5A). However, no DNA-bound dimer was observed. In contrast, holo-*Sa*BirA revealed free dimer and a species consistent with two holo-*Sa*BirA subunits in complex with DNA (Fig. 5B). Therefore, under high biotin conditions holo-*Sa*BirA can dimerise prior to binding DNA, such as occurs with *Ec*BirA. Like the wild-type protein, apo-*Sa*BirA-F123G was also competent to bind DNA as a single BirA subunit (Fig. 5C). Holo-*Sa*BirA-F123G was able to bind as a single subunit bound to DNA and well as a complex of two protein subunits bound to the mutated *SabioO* oligonucleotide (Fig. 5D). This datum suggests that once a single subunit of *Sa*BirA is bound, a second subunit can subsequently bind and stabilize the complex. This DNA-induced protein dimerization can proceed despite the inability of *Sa*BirA-F123G to dimerise in solution. EMSA analysis did not detect holo-*Sa*BirA binding to the mutated oligonucleotide containing only a single BirA binding site (Supporting Fig. S6). It is likely that the experimental conditions of an EMSA, whereby a protein:DNA complex must endure slow (minutes to hours) electrophoretic separation, may not favour the relatively short lived interactions that native nESI-MS can measure on the millisecond timescale. Thus, the native nESI-MS data strongly suggests a pathway where *Sa*BirA dimers can assemble on the DNA one subunit at a time.

Bioinformatic analysis of *S. aureus* BirA sequences. Several groups have now reported studies characterising BirA from *S. aureus*[24,25,27,29,34]. However, these studies have used genes encoding *birA* that have been cloned from different strains bearing different sequences making it difficult to directly compare the findings between studies. It was necessary to identify a consensus *Sa*BirA sequence as detailed studies on the *E. coli* homologue have demonstrated that substitution of certain amino acids can disrupt networks of long-range bonding interactions that can effect homodimerization, DNA binding and transcriptional repression despite not being localized to the dimer interface nor DNA binding domain (Fig. 6A)[18,19,21,35]. To identify the most conserved BirA homolog amongst *S. aureus*, BirA protein sequences were obtained using available genomic datum. A multiple sequence alignment was performed upon 26 non-redundant *Sa*BirA sequences (Supporting Fig. S7), including Mu50 (used in this study and in[24,29,34]), Newman (used in the experiments reported in[25,27]) and the well-studied NCTC 8325 strain. The alignment indicated that the sequence of Mu50 BirA is highly represented amongst the *S. aureus* genomes analysed. The Mu50 sequence differs from both the Newman and NCTC 8325 strains, specifically at positions 247 (R247 in Mu50, I247 in Newman and NCTC 8325) and 272 (I272 in Mu50, T272 in Newman and NCTC 8325) (Fig. 6B,C). Out of the 26 unique sequences analysed, 20 possessed arginine at 247 (conserved with Mu50) and 22 species had isoleucine at position 272 (conserved with Mu50). These residues are located within the central catalytic domain of *Sa*BirA, within α-helical structures (R247 in α-helix 8 and I272 in α-helix 9, Supporting Fig. S8). Neither residue is implicated in interactions with the substrates, biotin or ATP, nor are they located close to the dimerization interface. However, it is possible these residues may influence long range bonding interactions that can influence protein dimerization and DNA binding. Noteworthy is a tyrosine at position

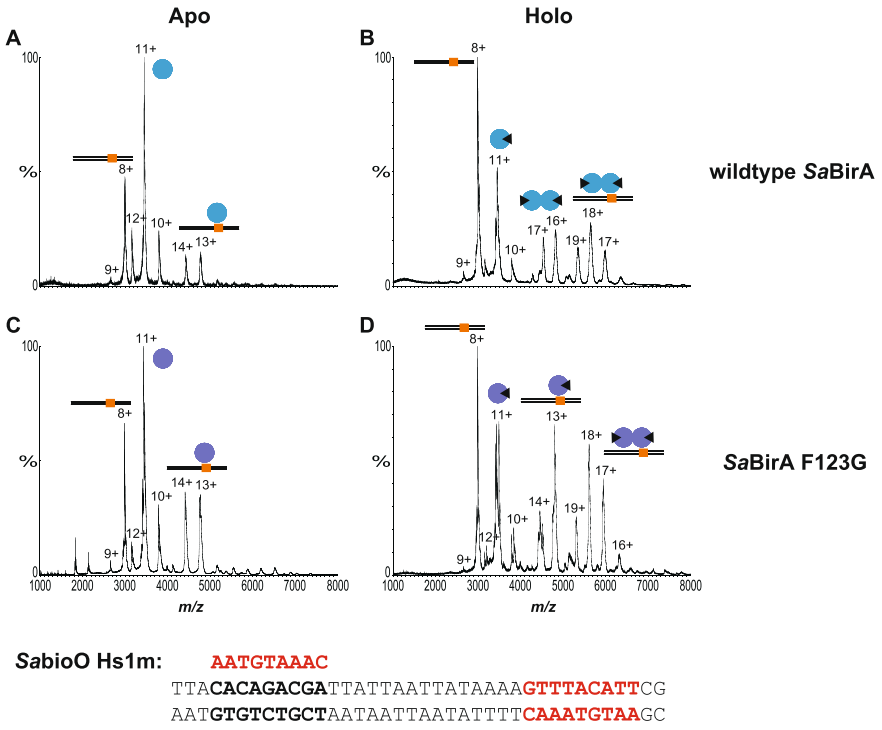


Figure 5. Native nano-electrospray ionization-mass spectrometry analysis of *Sa*BirA binding to DNA containing one binding site. nESI-MS spectra demonstrating the interaction between *SabioO* oligonucleotide containing only one BirA binding half site with (**A**) apo-*Sa*BirA, (**B**) holo-*Sa*BirA, (**C**) apo-*Sa*BirA-F123G and (**D**) holo-*Sa*BirA-F123G. Monomers are denoted by a single sphere, homodimers are shown by two joined-spheres and the presence of biotinyl-5′-AMP is represented with a black triangle. Black lines represent double stranded oligonucleotides containing one BirA binding site (orange). The charged states of the major species are annotated on the spectra. The DNA sequences of the oligonucleotides used in the assay are shown below with the wild-type half sites highlighted in red and the mutated half site in bold text.

182 that was conserved in 25 strains including Mu50 and Newman, but which is a phenylalanine in NCTC 8325. Recent studies have reported that substitution of the equivalent tyrosine in *Ec*BirA (Y178) disrupts an electrostatic network of binding interactions that alters protein dimerization[18,19]. As the Mu50 sequence employed here is the most highly representative homologue amongst *S. aureus*, this sequence should be considered the prototypical *Sa*BirA in future experiments.

Discussion

In certain bacteria BirA is a bifunctional protein that is both a biotin protein ligase and transcriptional repressor of biotin biosynthesis. The biosynthesis of biotin is metabolically costly, requiring the activity of at least four gene products and 20 ATP equivalents per biotin molecule[36]. Therefore, the activity of BirA allows cells to find the necessary balance between cellular demand, supply from the environment and *de novo* synthesis. Here we study the BirA protein from two clinically important bacteria, namely *E. coli* and *S. aureus*. *E. coli* is considered the prototypical bacteria, and its BirA protein has been the subject of many biochemical, structural and genetic studies. The *S. aureus* homologue has not been investigated in as much detail. The two bacteria colonize different micro niches where the availability of environmental biotin is distinctive. *E. coli* is a part of the intestinal microflora which contributes to the synthesis and release of substantial amounts of biotin. Recent mouse studies estimate the concentration of biotin in the ileum to be 450–700 nM[37]. Conversely, *S. aureus* possess the extraordinary ability to colonise a wide range of niche microenvironments and is responsible for various disorders affecting the skin, respiratory organs, soft tissues, bones and joints. The bioavailability of biotin at these sites is lower than in the intestine. Accordingly, we propose that *S. aureus* BirA has evolved unique properties that allow the bacteria to adapt to low biotin environments. In this study, we performed detailed *in vivo* and *in vitro* analyses to obtain new insights into the transcription regulation mechanisms of *Sa*BirA. We utilized the *Sa*BirA from the Mu50 methicillin and vancomycin resistant strain of *S. aureus*[38]. Our bioinformatics analysis has demonstrated this particular variant of *Sa*BirA represents the most conserved *Sa*BirA sequence among *S. aureus* strains. Furthermore, Mu50 *Sa*BirA has been the target for the discovery of new classes of antibiotics designed to treat MRSA infections[23,39–41]. We also performed our study alongside the well-characterized BirA from *E. coli*, in order to highlight the differences in dimerization and DNA-binding mechanisms between the two homologues.

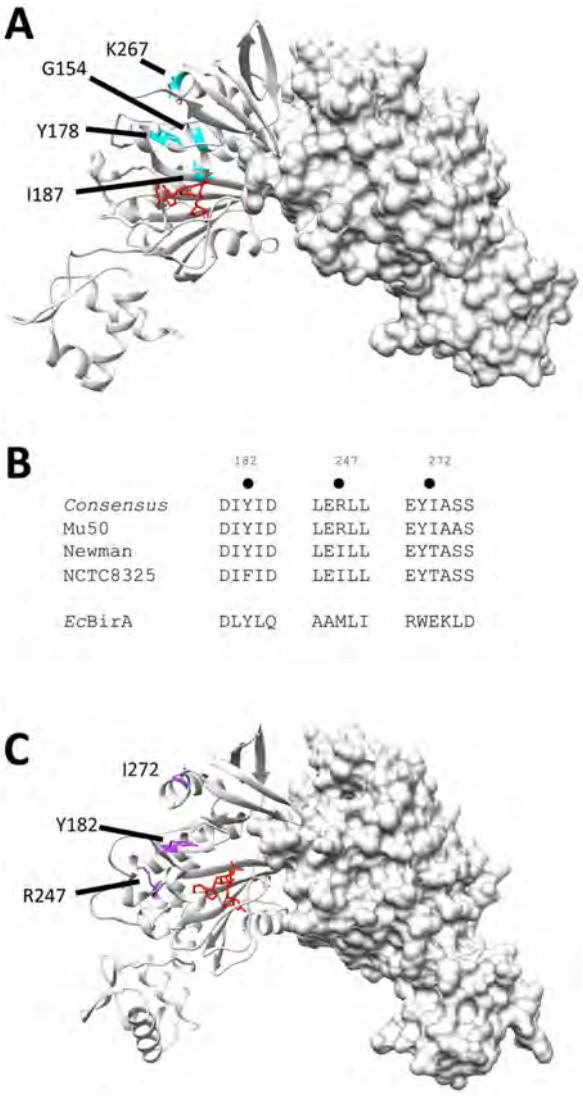


Figure 6. *Sa*BirA sequence analysis. (**A**) Crystal structure of *Ec*BirA [PDBID 2EWN[6]] highlighting amino acids that influence homodimerization when mutated. (**B**) 26 non-redundant *Sa*BirA sequences were aligned to determine the consensus sequence (see Supporting Fig. S7 for full alignment). Differences between *S. aureus* strains Mu50, Newman and NCTC 8325 are shown here. The *Ec*BirA sequence is also aligned (below) for comparison. C. Crystal structure of *Sa*BirA [PDBID 4DQ2[23]] showing the position of amino acids that are not conserved between *S. aureus* strains Mu50, Newman and NCTC8325.

Previous studies performed on Mu50 *Sa*BirA revealed weak dimerization of the apo-form of this protein ($K_D^{2\text{-}1} = 29 \pm 0.2\,\mu M$)[29] suggesting that in its apo-state, this protein is likely to be monomeric within the intracellular environment. Hence, we proposed that *Sa*BirA-F123G, as a dimerization-deficient mutant in solution, would mimic the oligomeric state of this apo-*Sa*BirA *in vivo*. Whilst the analogous *E. coli* mutant was clearly devoid of DNA-binding activity *in vivo*, as reported in the literature[31] and seen here, our results clearly demonstrated *Sa*BirA-F123G functioned as an effective transcriptional repressor for all three *S. aureus* target promoters *in vivo*. Furthermore, even at environmental biotin concentrations as low as 1 nM, wild-type *Sa*BirA and *Sa*BirA-F123G demonstrated some repression compared to the no repressor control. This phenomenon was in contrast to *Ec*BirA. This suggests that *Sa*BirA can either respond to lower concentrations of environmental biotin (unlikely since K_M values for biotin are similar ($0.3 \pm 0.1\,\mu M$ for *Ec*BirA[42] v $1.01 \pm 0.16\,\mu M$ for *Sa*BirA[23]), or that *Sa*BirA partially occupies the DNA even when biotin is limiting, and therefore, the protein is in the apo-state.

In vitro analysis supported the contention that the intrinsically monomeric *Sa*BirA-F123G or wild-type apo-*Sa*BirA can bind DNA, enabling *in vivo* repression. The Mu50 strain wild-type *Sa*BirA used here was unable to dimerize in solution at a protein concentration of 10 µM in its apo-form, as indicated by native nESI-MS. This

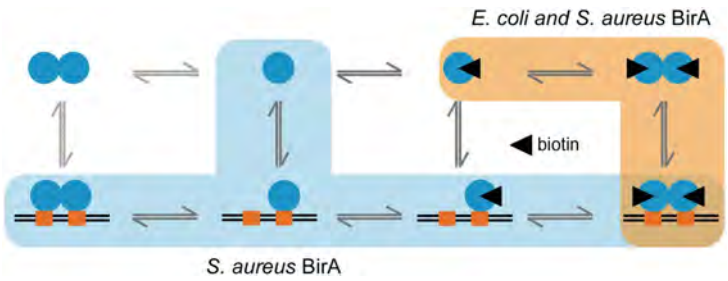


Figure 7. Proposed DNA binding pathway of *Sa*BirA. In biotin-limiting conditions, BirA cannot dimerise in solution. However, *Sa*BirA monomers (blue ovals) can bind to individual half sites (orange boxes), with the interaction presumably stabilized by interaction between the monomers on the DNA. Under high biotin conditions (biotin is indicated by a black triangle), *Sa*BirA follows the same binding mechanism as holo-*Ec*BirA, where a pre-formed dimer binds to DNA (light orange shading).

finding agrees with our previous work which has demonstrated the apo-protein to be monomeric by small angle X-ray scattering assay[29] and crystallises as a single subunit[24]. It also agrees with recent work conducted on *Sa*BirA from the Newman strain reported by Wang and Beckett[25] where analytical ultracentrifugation (AUC) failed to detect homodimerization. This is in contrast to our previous AUC datum where dimerisation of apo-*Sa*BirA from Mu50 was observed[29]. It should be noted that different protein sequences were employed between these two studies (discussed further below). Additionally, this may have been the result of contaminating co-repressor, biotinyl-5′-AMP, that was not detected by the Streptavidin blot method used to assay the purified protein[27]. Here, a superior technique of native nESI-MS has been utilized to unambiguously detect not only the oligomeric state of the protein, but to simultaneously determine whether these species are truly devoid of bound ligands. This is the first report where native mass spectrometry has been employed to study BirA. No dimeric form of *Sa*BirA-F123G was observed in either the apo or holo forms in solution, supporting its characterization as a monomeric mutant.

Furthermore, our results suggest dimerization of apo-*Sa*BirA and *Sa*BirA-F123G is mediated by DNA binding, as native nESI-MS detected species with molecular weights equivalent to two monomers of *Sa*BirA in complex with the *SabioY* recognition sequence oligonucleotide. Likewise, nESI-MS revealed wild-type and *Sa*BirA-F123G bound a DNA probe containing only a single half site in both the apo and holo-states but with different stoichiometry equivalents. Whilst the apo-proteins only bind as one subunit, biotinyl-5′-AMP allows the dimerization of wild-type *Sa*BirA prior to DNA binding but also allowed *Sa*BirA-F123G to dimerise on the DNA, despite not detectably dimerising in solution. The binding of *Sa*BirA-F123G to DNA agrees with previous studies on apo-*Sa*BirA where the K_D for DNA binding was 6-fold lower compared to holo-*Sa*BirA as measured by EMSA[29], and 60-fold lower as measured by fluorescence anisotropy[27].

By combining our *in vitro* and *in vivo* data, we propose two possible pathways for *Sa*BirA to assemble on DNA, as illustrated in Fig. 7. In high biotin conditions, protein dimerization occurs in solution and precedes DNA binding, as has been well documented for *Ec*BirA[5,7]. The pre-formed homodimer orients the two helix-turn-helix motifs present in the N-terminus of the protein such that they can simultaneously occupy both half sites on the target operator. Here we present evidence that binding to DNA can also proceed via the sequential assembly of two monomers onto the DNA, a pathway which is much less preferred for *Ec*BirA. This alternative pathway is likely be the basis for the lower biotin threshold required to initiate the transcriptional repression observed with *Sa*BirA compared to *Ec*BirA, as highlighted in a previous study[26] and seen with the partial occupation of *Sa*BirA on DNA causing repression at low biotin concentrations (1 nM) in the *in vivo* reporter assay[26]. Whilst the DNA mediated dimerization pathway has been observed with the *Sa*BirA from the Mu50 strain, which we have identified here as the prototypical example in *S. aureus*, it remains to be seen whether BirA from other species behave in a similar manner. The findings from this study may provide new insights into how biotin-regulated gene expression occurs in *S. aureus* allowing it to adapt to the low biotin environments it colonises.

Our bioinformatics analysis has identified multiple polymorphisms amongst *S. aureus* strains which may give rise to BirA variants with different homo-dimerization and DNA binding activities. Mutagenesis studies that target specific amino acids in *Ec*BirA, together with structural biology and biophysical assays, are helping to understand mechanisms of allosteric communication between distant sites in the protein and the impact upon their function[18,21,35]. Noteworthy is mutation of tyrosine 178 to cysteine (equivalent to Y182 in *S. aureus* Mu50, and F182 in NCTC 8325) that results in a super-repressor phenotype in whole cell reporter assays of transcription[18]. This amino acid does not reside directly in the dimer interface but, instead, contributes to complex network of electrostatic interactions that are proposed to influence the structural alignment of the central (ligand-binding) and N-terminal (DNA-binding) domains of BirA that effect homo-dimerization and DNA binding[19]. Examination of available crystal structures of the Mu50 *Sa*BirA[23,24,29] likewise revealed that Y182 contributes to a network of electrostatic interactions involving K176 (that resides in a conserved KWPND motif in the active site) and D322 that resides in the DNA binding domain. Our analysis of genomic information also identified two differences in the BirA proteins from Mu50 and Newman at positions 247 and 272. Mutational studies, such as those performed to characterise the *Ec*BirA mutants[18,19], are required to probe the structure-function relationships in more detail. We propose that the *Sa*BirA variants identified here present a useful system with which to further probe long distance modulation of protein allostery.

Methods

Materials. *E. coli* strain JD26186 (*bioC*::Kan), based on host strain KP7600, is a transposon-disrupted bioC biotin auxotroph, and was obtained from the National BioResource Project (NIG, Japan). Bacteria were cultured at 37 °C with vigorous shaking in Luria Bertani media containing the appropriate antibiotic. Plasmid extractions were performed using the Plasmid QIAprep Spin Miniprep Kit (Qiagen) and genomic extractions were performed using the Wizard® Genomic DNA purification kit (Promega). All molecular biology enzymes (DNA polymerase, ligases and restriction enzymes) and buffers were supplied by New England Biolabs. Oligonucleotides, purchased from Geneworks Ptd Ltd, are shown in Supporting Table S1.

Nucleic acid manipulations. All plasmids employed in this study are shown in Supporting Table S2. Site directed mutagenesis (Quikchange Site Directed Mutagenesis kit, Stratagene), was performed upon plasmid pGEM-*Ec*BirA-H6[29] using oligonucleotides B460 and B461, yielding pGEM-*Ec*BirA-R119W-H6. The *EcbirA-R119W-H6* coding region was subsequently cloned into either pET16b (Novagen) for recombinant protein expression, or the integration vector pIT4_TL_152002[43], yielding pIT4_TL_*Ec*BirA-R119W-H6. Similarly, the *SabirA-F123G-H6* coding region was excised from pGEM-*Sa*BirA-F123G-H6[29] and ligated into pIT4_TL_152002, yielding pIT4_TL_*Sa*BirA-F123G-H6.

Generation of bacterial strains for *in vivo* reporter assays. All strains of *E. coli* employed in this study are shown in Supporting Table S3. β-galactosidase reporter assays were performed on *E. coli* reporter strains generated in[26]. Both the β-galactosidase reporters and the BirA expression modules were site-specifically integrated in the chromosome in single copy (Fig. 1A). BirA binding sequence was cloned into pLacZ_SHTrim plasmid upstream of the β-galactosidase reporter gene. BirA expression (*Ec*BirA, *Sa*BirA and their dimerization defective mutants *Ec*BirA-R119W and *Sa*BirA-F123G) was under control of the placUV5 promoter. These were constructed using the methods outlined in[26,44]. The β-galactosidase reporter assays were conducted as described[43,45].

Protein methods. Recombinant expression of all BirA proteins employed in this study, as well as their purification by immobilized metal ion affinity chromatography (IMAC), were performed essentially as previously described[29,34]. However, the following modifications were included to ensure removal of biotinyl-5′-AMP from the protein preparations. Prior to IMAC purification, the cell lysate was incubated at 4 °C for a minimum of 1 hour with 10 µL 50% slurry Streptavidin-sepharose High Performance (GE Healthcare Life Sciences) per millilitre of lysate. This was centrifuged at 4000 rpm, 4 °C for 10 minutes, to remove the resin. The cleared lysate was then incubated for 1 hour at 37 °C with 6 mg biotin-accepting substrate GST-*Sa*PC90 per 10 ml of lysate prior to IMAC. Purified proteins were dialyzed overnight in 4 L of storage buffer (50 mM Tris pH 8.0, 100 mM KCl, 1 mM DTT, 5% (v/v) glycerol) and stored at −80 °C until required. Protein SDS-PAGE was performed using NuPage™ Bis-Tris 4–12% gels (Invitrogen).

Native nano-electrospray ionisation-mass spectrometry (nESI-MS). Purified BirA proteins were buffer exchanged into 200 mM ammonium acetate pH 6.85 (Sigma) using Amicon Ultra-0.5 MWCO 10,000 centrifugal filter units (Merck Millipore). Holo-enzyme samples (at least 100 µM, except for *Ec*BirA-R119W that was 30 µM) were prepared by pre-incubating the apo-proteins with 500 µM biotin, 1 mM ATP and 1 mM $MgCl_2$ on ice for at least 1 hour prior to buffer exchange. HPLC-purified and annealed, double stranded oligonucleotides containing the operator sequences of interest were purchased from Integrated DNA Technology (USA). These were desalted into 200 mM ammonium acetate using Illustra MicroSpin G-25 columns (GE Healthcare). Oligonucleotide concentrations were quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific). Proteins were quantified following buffer exchange via Bradford assay, and then diluted to 10 µM in 200 mM ammonium acetate for analysis by nESI-MS. Protein:DNA complexes were analyzed by nESI-MS after diluting the oligonucleotide and protein in 200 mM ammonium acetate pH 6.9 to 5 µM and 10 µM respectively, and incubating at 4 °C for at least 1 hour prior to MS analysis.

Mass spectrometry (MS) measurements were performed on a Synapt HDMS system (Waters, UK) with the sample introduced by nano-electrospray ionization in the positive ion MS mode from platinum-coated borosilicate capillaries prepared in-house. Instrument parameters were optimized to remove adducts while preserving non-covalent interactions, and were as follows; capillary voltage, 1.5 kV; cone voltage, 60 V; trap collision energy, 20 V; transfer collision energy, 15 V; source temperature, 50 °C; backing pressure, 3.95 mbar. Data analysis was performed using manual peak finding in MassLynx software (version 4.1).

Electrophoretic mobility shift assay (EMSA). Binding reactions were performed at room temperature for 30 minutes using EMSA buffer (50 mM Tris pH 8.0, 50 mM NaCl, 0.1 mM biotin, 1 mM ATP, 1 mM $MgCl_2$ and 10% (v/v) glycerol) together with 10 nM of the double stranded oligonucleotide (HPLC-purified as utilized for nESI-MS) and varying concentrations of BirA. Fractionation of the samples was performed using 4–12% TBE polyacrylamide gels (Life Technologies) run at 100 volts (constant) for 45 minutes and stained in GelRed (Biotium) solution for 5 minutes. After washing five times in distilled water, the gels were imaged using a ChemiDoc imager (Bio-Rad). Each EMSA experiment was performed in triplicate.

***S. aureus* BirA multiple sequence alignment.** BirA sequences of *S. aureus* were compared as follows: the protein cluster PCLA_885364 (accessed on 15/5/17) containing 71 entries (which encompasses all non-redundant proteins identified as biotin-acetyl-CoA-carboxylase ligase in the Genus Staphylococcus) was obtained from NCBI[46,47] and further filtered for the clade *Staphylococcus aureus* 19988 to make a list of 26 unique protein sequences. These sequences were then aligned using Clustal Omega[48]. The alignment was then compared using UGENE[49] and the position of the polymorphisms were mapped onto the *Sa*BirA structure (PDB accession

4DQ2[23]) using UCSF Chimera software (version 1.12). The *Ec*BirA sequence for comparison was obtained from the PDB accession 2EWN[6].

References

1. Satiaputra, J., Shearwin, K. E., Booker, G. W. & Polyak, S. W. Mechanisms of biotin-regulated gene expression in microbes. *Synth Syst Biotechnol* **1**, 17–24, https://doi.org/10.1016/j.synbio.2016.01.005 (2016).
2. Cronan, J. E. Biotin and lipoic acid: synthesis, attachment, and regulation. *EcoSal Plus* **6**, https://doi.org/10.1128/ecosalplus.ESP-0001-2012 (2014).
3. Zhao, H., Streaker, E., Pan, W. & Beckett, D. Protein-protein interactions dominate the assembly thermodynamics of a transcription repression complex. *Biochemistry* **46**, 13667–13676, https://doi.org/10.1021/bi7013097 (2007).
4. Streaker, E. D. & Beckett, D. Coupling of site-specific DNA binding to protein dimerization in assembly of the biotin repressor-biotin operator complex. *Biochemistry* **37**, 3210–3219, https://doi.org/10.1021/bi9715019 (1998).
5. Streaker, E. D. & Beckett, D. Coupling of protein assembly and DNA binding: biotin repressor dimerization precedes biotin operator binding. *J Mol Biol* **325**, 937–948 (2003).
6. Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D. & Matthews, B. W. Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution. *J Mol Biol* **357**, 509–523, https://doi.org/10.1016/j.jmb.2005.12.066 (2006).
7. Adikaram, P. R. & Beckett, D. Protein:protein interactions in control of a transcriptional switch. *J Mol Biol* **425**, 4584–4594, https://doi.org/10.1016/j.jmb.2013.07.029 (2013).
8. Adikaram, P. R. & Beckett, D. Functional versatility of a single protein surface in two protein:protein interactions. *J Mol Biol* **419**, 223–233, https://doi.org/10.1016/j.jmb.2012.03.010 (2012).
9. Xu, Y. & Beckett, D. Evidence for interdomain interaction in the *Escherichia coli* repressor of biotin biosynthesis from studies of an N-terminal domain deletion mutant. *Biochemistry* **35**, 1783–1792, https://doi.org/10.1021/bi952269e (1996).
10. Barker, D. F. & Campbell, A. M. Genetic and biochemical characterization of the birA gene and its product: evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in *Escherichia coli*. *J Mol Biol* **146**, 469–492 (1981).
11. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc Natl Acad Sci USA* **89**, 9257–9261 (1992).
12. Streaker, E. D. & Beckett, D. Ligand-linked structural changes in the *Escherichia coli* biotin repressor: the significance of surface loops for binding and allostery. *J Mol Biol* **292**, 619–632, https://doi.org/10.1006/jmbi.1999.3086 (1999).
13. Cronan, J. E. The *E. coli* bio operon: transcriptional repression by an essential protein modification enzyme. *Cell* **58**, 427–429 (1989).
14. Streaker, E. D., Gupta, A. & Beckett, D. The biotin repressor: thermodynamic coupling of corepressor binding, protein assembly, and sequence-specific DNA binding. *Biochemistry* **41**, 14263–14271 (2002).
15. Zhao, H., Naganathan, S. & Beckett, D. Thermodynamic and structural investigation of bispecificity in protein-protein interactions. *J Mol Biol* **389**, 336–348, https://doi.org/10.1016/j.jmb.2009.04.009 (2009).
16. Eisenstein, E. & Beckett, D. Dimerization of the *Escherichia coli* biotin repressor: corepressor function in protein assembly. *Biochemistry* **38**, 13077–13084 (1999).
17. Kwon, K., Streaker, E. D., Ruparelia, S. & Beckett, D. Multiple disordered loops function in corepressor-induced dimerization of the biotin repressor. *J Mol Biol* **304**, 821–833, https://doi.org/10.1006/jmbi.2000.4249 (2000).
18. Chakravartty, V. & Cronan, J. E. Altered regulation of *Escherichia coli* biotin biosynthesis in BirA superrepressor mutant strains. *J Bacteriol* **194**, 1113–1126, https://doi.org/10.1128/jb.06549-11 (2012).
19. He, C., Custer, G., Wang, J., Matysiak, S. & Beckett, D. Superrepression through altered corepressor-activated protein:protein interactions. *Biochemistry* **57**, 1119–1129, https://doi.org/10.1021/acs.biochem.7b01122 (2018).
20. Buoncristiani, M. R., Howard, P. K. & Otsuka, A. J. DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (BirA) of *Escherichia coli*. *Gene* **44**, 255–261 (1986).
21. Eginton, C., Cressman, W. J., Bachas, S., Wade, H. & Beckett, D. Allosteric coupling via distant disorder-to-order transitions. *J Mol Biol* **427**, 1695–1704, https://doi.org/10.1016/j.jmb.2015.02.021 (2015).
22. Barker, D. F. & Campbell, A. M. Use of bio-lac fusion strains to study regulation of biotin biosynthesis in. *Escherichia coli. J Bacteriol* **143**, 789–800 (1980).
23. Soares da Costa, T. P. *et al*. Selective inhibition of Biotin protein ligase from *Staphylococcus aureus*. *J. Biol. Chem.* **287**, 17823–17832, https://doi.org/10.1074/jbc.M112.356576 (2012).
24. Pendini, N. R. *et al*. Structural characterization of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci.* **22**, 762–773, https://doi.org/10.1002/pro.2262 (2013).
25. Wang, J. & Beckett, D. A conserved regulatory mechanism in bifunctional biotin protein ligases. *Protein Sci* **26**, 1564–1573, https://doi.org/10.1002/pro.3182 (2017).
26. Satiaputra, J. *et al*. Biotin-mediated growth and gene expression in Staphylococcus aureus is highly responsive to environmental biotin. *Appl Microbiol Biotechnol* **102**, 3793–3803, https://doi.org/10.1007/s00253-018-8866-z (2018).
27. Henke, S. K. & Cronan, J. E. The *Staphylococcus aureus* group II biotin protein ligase BirA is an effective regulator of biotin operon transcription and requires the DNA binding domain for full enzymatic activity. *Mol Microbiol* **102**, 417–429, https://doi.org/10.1111/mmi.13470 (2016).
28. Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. *Genome Res* **12**, 1507–1516, https://doi.org/10.1101/gr.314502 (2002).
29. Soares da Costa, T. P. *et al*. Dual roles of F123 in protein homodimerization and inhibitor binding to biotin protein ligase from *Staphylococcus aureus*. *Mol. Microbiol.* **91**, 110–120, https://doi.org/10.1111/mmi.12446 (2014).
30. Hanson, C. L. & Robinson, C. V. Protein-nucleic acid interactions and the expanding role of mass spectrometry. *J Biol Chem* **279**, 24907–24910, https://doi.org/10.1074/jbc.R300037200 (2004).
31. Kwon, K. & Beckett, D. Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci* **9**, 1530–1539, https://doi.org/10.1110/ps.9.8.1530 (2000).
32. Taniguchi, Y. *et al*. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538, https://doi.org/10.1126/science.1188308 (2010).
33. Priest, D. G. *et al*. Quantitation of the DNA tethering effect in long-range DNA looping *in vivo* and *in vitro* using the Lac and lambda repressors. *Proc Natl Acad Sci USA* **111**, 349–354, https://doi.org/10.1073/pnas.1317817111 (2014).
34. Pendini, N. R., Polyak, S. W., Booker, G. W., Wallace, J. C. & Wilce, M. C. J. Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from *Staphylococcus aureus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64**, 520–523, https://doi.org/10.1107/S1744309108012244 (2008).
35. Wang, J., Custer, G., Beckett, D. & Matysiak, S. Long distance modulation of disorder-to-order transitions in protein allostery. *Biochemistry* **56**, 4478–4488, https://doi.org/10.1021/acs.biochem.7b00496 (2017).
36. Feng, Y. *et al*. A Francisella virulence factor catalyses an essential reaction of biotin synthesis. *Mol Microbiol* **91**, 300–314, https://doi.org/10.1111/mmi.12460 (2014).
37. Yang, B., Feng, L., Wang, F. & Wang, L. Enterohemorrhagic *Escherichia coli* senses low biotin status in the large intestine for colonization and infection. *Nat Commun* **6**, 6592, https://doi.org/10.1038/ncomms7592 (2015).

38. Kuroda, M. *et al.* Whole genome-sequencing of methicillin-resistant *Staphylococcus aureus*. *The Lancet* **357**, 1225–1240 (2001).
39. Tieu, W. *et al.* Improved synthesis of biotinol-5′-AMP: Implications for antibacterial discovery. *ACS Med. Chem. Lett.* **6**, 216–220, https://doi.org/10.1021/ml5004750 (2015).
40. Tieu, W. *et al.* Heterocyclic acyl-phosphate bioisostere-based inhibitors of *Staphylococcus aureus* biotin protein ligase. *Bioorg. Med. Chem. Lett.* **24**, 4689–4693, https://doi.org/10.1016/j.bmcl.2014.08.030 (2014).
41. Paparella, A. S. *et al.* Halogenation of Biotin Protein Ligase Inhibitors Improves Whole Cell Activity against Staphylococcus aureus. *ACS Infect Dis* **4**, 175–184, https://doi.org/10.1021/acsinfecdis.7b00134 (2018).
42. Chapman-Smith, A. *et al.* The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity. *Protein Sci* **10**, 2608–2617, https://doi.org/10.1110/ps.22401 (2001).
43. St-Pierre, F. *et al.* One-step cloning and chromosomal integration of DNA. *ACS Synthet. Biol.* **2**, 537–541, https://doi.org/10.1021/s400021j (2013).
44. Cui, L., Murchland, I., Dodd, I. B. & Shearwin, K. E. Bacteriophage lambda repressor mediates the formation of a complex enhancer-like structure. *Transcription* **4**, 201–205 (2013).
45. Cui, L. & Shearwin, K. E. Clonetegration Using OSIP Plasmids: One-Step DNA Assembly and Site-Specific Genomic Integration in Bacteria. *Methods Mol Biol* **1472**, 139–155, https://doi.org/10.1007/978-1-4939-6343-0_11 (2017).
46. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**, D733–745 (2016).
47. Sayers, E. W. *et al.* Database resources of the National Centre for Biotechnology Information. *Nucleic Acids Res* **37**, D5–D15 (2009).
48. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7** (2011).
49. Okonechnikov, K., Golosova, O. & Fursov, M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).

Acknowledgements

We thank the National BioResource Project (NIG, Japan) for the provision of bacterial strains. We would like to thank Dr. Ian Dodd and Dr. Andrew Hao for their assistance and useful discussions, as well as Mr. James Merritt and Ms. Jade Foeng for their assistance with assays. This work was supported by the National Health and Medical Research Council of Australia [Project Grant APP1068885 to SWP and GWB], and the Australian Research Council [Discovery Project DP160101450 to KES]. We are also grateful to the Wallace and Carthew families for their financial support. JS was the recipient of a University of Adelaide Faculty Divisional Scholarship, LMS was the recipient of an Australian Postgraduate Award and AJH was the recipient of a University of Adelaide Scholarship.

Author Contributions

S.W.P. and G.W.B. designed the study, J.S., L.M.S., A.J.H. and T.L.P. acquired data, T.L.P., K.E.S. and S.W.P. analysed and interpreted the data, and J.S., L.M.S., K.E.S. and S.W.P. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-39398-6.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supporting Information

Native mass spectrometry identifies an alternative DNA-binding pathway for BirA from *Staphylococcus aureus*

Jiulia Satiaputra[1#a], Louise M. Sternicki[1#], Andrew J. Hayes[1b], Tara. L. Pukala[2], Grant W. Booker[1], Keith E. Shearwin[1] and Steven W. Polyak[1*c]

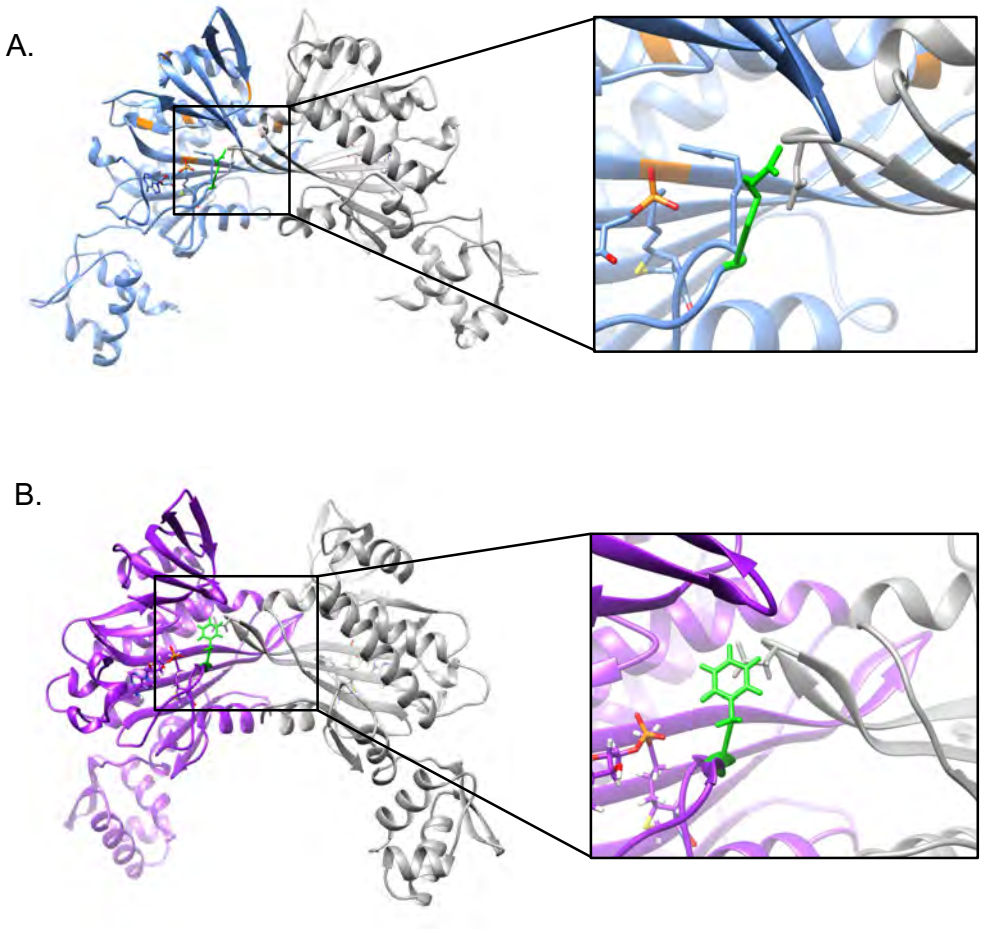
[1] School of Biological Sciences, University of Adelaide, South Australia 5005, Australia;
[2] School of Physical Sciences, University of Adelaide, South Australia 5005, Australia

Present Address: [a] Harry Perkins Institute of Medical Research, Western Australia, 6008, Australia. [b] Faculty of Health and Medical Sciences, Adelaide, South Australia, 5005, Australia. [c] School of Pharmacy and Medical Sciences, University of South Australia, South Australia 5001.

* To whom correspondence should be addressed. Tel: +61 8 8313 4062; Email: steven.polyak@adelaide.edu.au

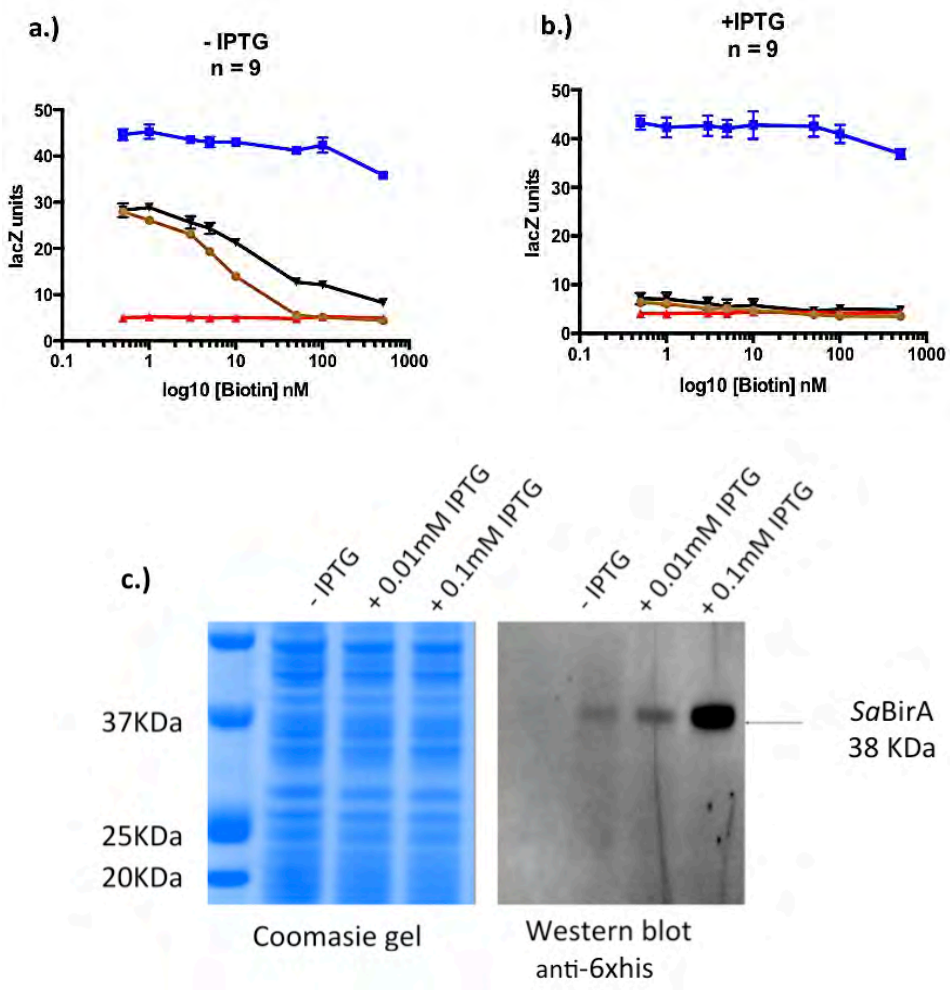
#J.S. and L.M.S. should be considered joint first author.

Supporting Figures



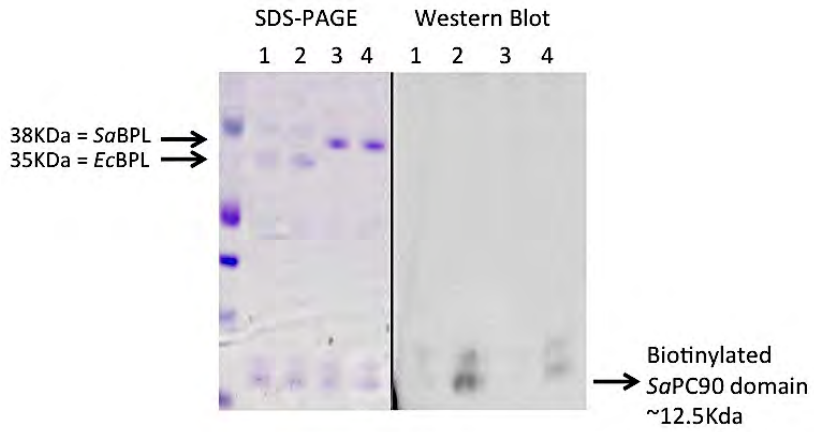
Supporting Figure S1. The dimeric structures of (A) *Ec*BirA and (B) *Sa*BirA reveal the equivalent residues are mutated to make the dimerization-impaired mutants *Ec*BirA-R119W and *Sa*BirA-F123G, respectively.

*Ec*BirA (PDB 2EWN) [1] is shown with one monomer in blue and the other in grey, whilst *Sa*BirA (PDB 4DQ2) [2] has one monomer shown in purple and the other in grey. The residues that are mutated to produce the monomeric mutants are shown in green, whilst the residues with which they interact in the partner subunit are in grey. Both *Ec*BirA and *Sa*BirA have the co-repressor (and reaction intermediate) analogue, biotinol-5′-AMP, bound.



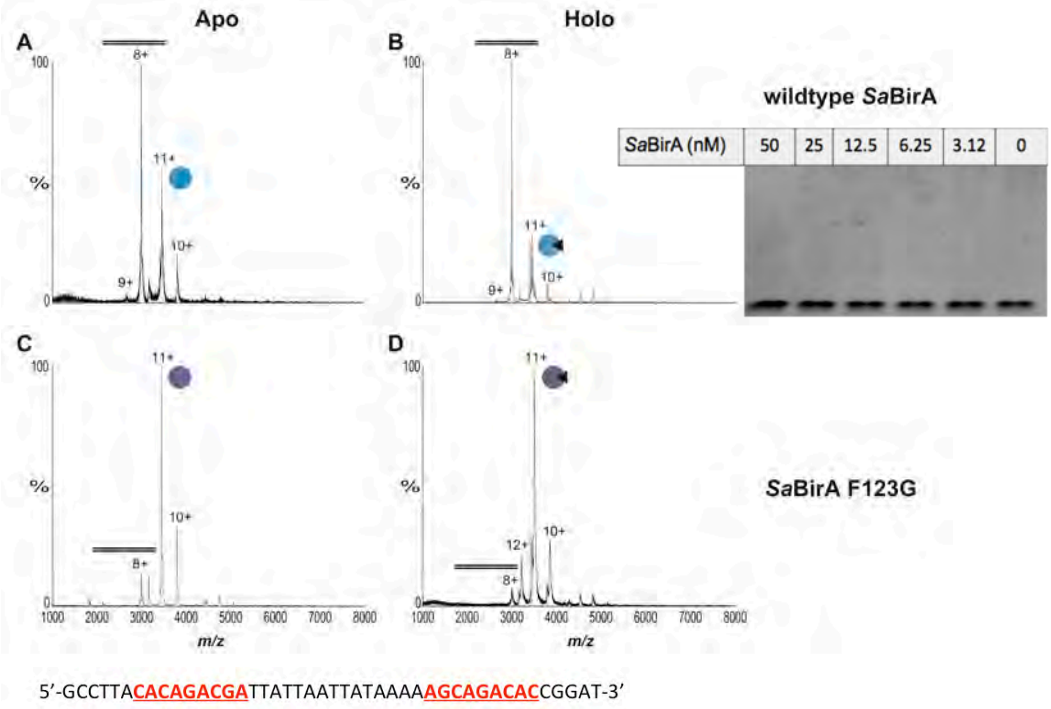
Supporting Figure S2. Repressor activity in the presence and absence of IPTG.

In vivo repressor activity of *Sa*BirA (blue curves) and *Sa*BirA F123G (brown curves) in either a) the absence of IPTG or (b) 0.01mM IPTG. The blue curves represent no-repressor controls and red curves represent no-promoter controls. β-galactosidase activity was measured in response to varying concentrations of biotin in the growth media. Error bars denote S.E.M. of n = 9. (c.) IPTG inducible expression of *Sa*BirA-H6 was detected by Western blot probed with anti-H6 antibody (right panel). A Coomassie blue stained gel was included as a control for protein loading (left panel).



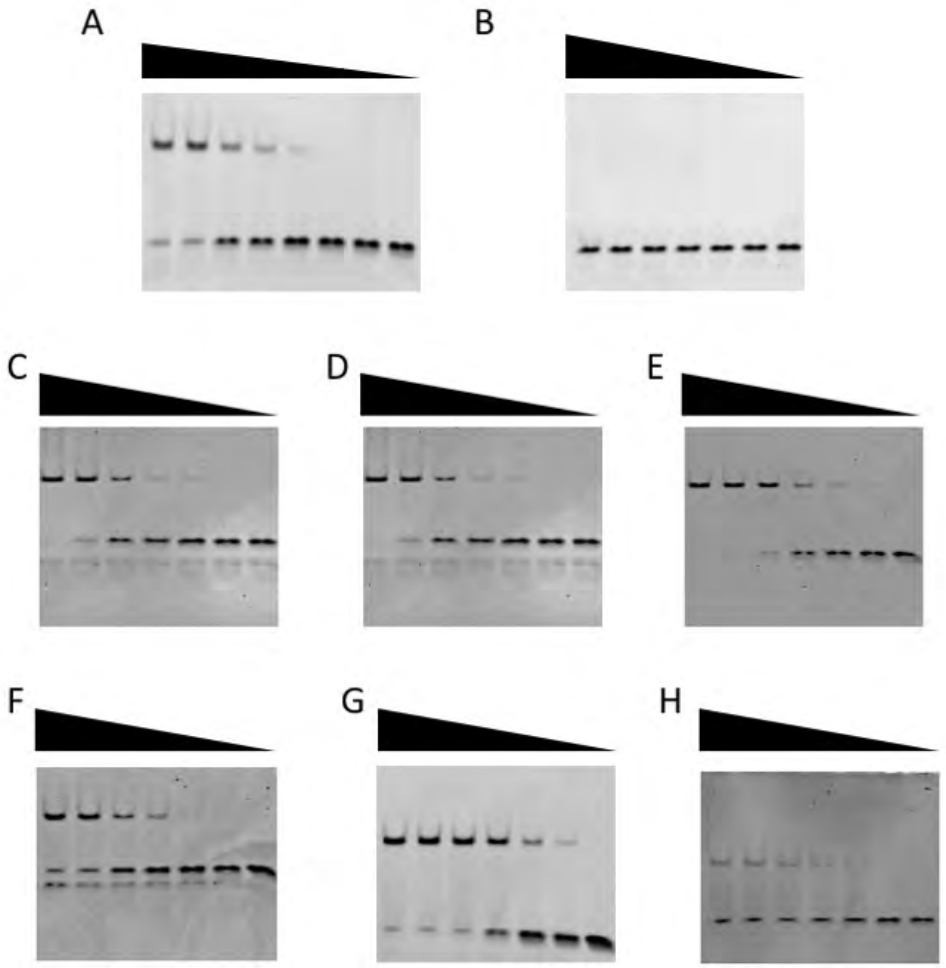
Supporting Figure S3. Biotinyl transferase assay.

Streptavidin-blot analysis was performed upon the products of an *in vitro* biotinylation assay to confirm biotinyl-5′-AMP did not co-purify with apo-*SaBirA* or apo-*Ec*BPL, as previously described [3]. The assay used SaPC90 as the biotin-accepting protein substrate. Reactions containing either (1) apo-*Ec*BPL, (2) holo-*Ec*BPL, (3) apo-*SaBirA* or (4) holo-*SaBirA* were investigated. The absence of bands on the Streptavidin blot in lanes 1 and 3 suggested biotinyl-5′-AMP had not been co-purified with the BirAs. The corresponding SDS-PAGE analysis is presented as a loading control.



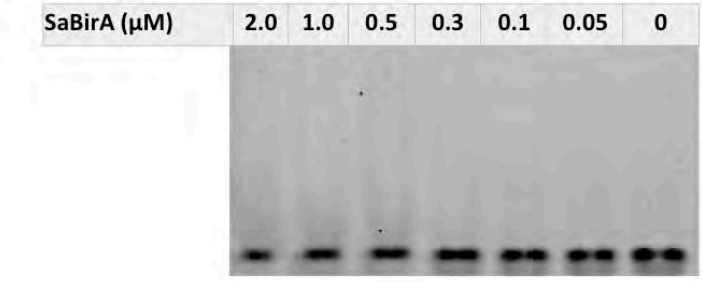
Supporting Figure S4. Control for non-specific binding between *Sa*BirA and a mutated *SabioO* oligonucleotide containing no BirA binding sites.

A protein:DNA complex was not detected by native nESI-MS using a mutated oligonucleotide and (A) apo-*Sa*BirA, (B) holo-*Sa*BirA, (C) apo-*Sa*BirA-F123G or (D) holo-*Sa*BirA-F123G. The DNA sequence is shown with the mutated half site sequences highlighted in red. Similarly, EMSA analysis failed to detect protein binding (insert). Within the MS spectra, double black lines represent the oligonucleotide, light blue spheres signify *Sa*BirA and dark purples are *Sa*BirA-F123G. Black triangles signify the presence of biotinyl-5′-AMP and, therefore, holo-protein. Measured masses of all the products are shown in Supporting Table S6.

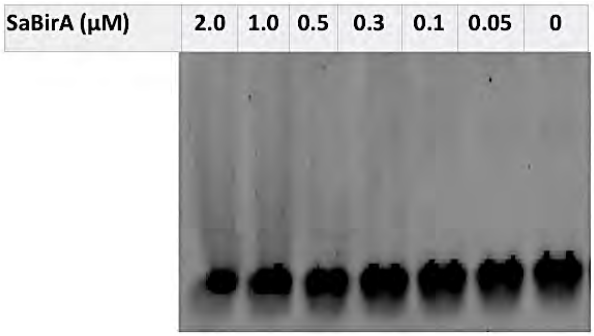


Supporting Figure S5. Full length blots from cropped blots in Figure 4.

EMSA of the interaction between *EcbioO* and (A) *Ec*BirA or (B) *Ec*BirA-R119W. Also shown is *Sa*BirA binding to (C) *SabioO*, (D) *bioY* or (E) *yhfS-yhfT* and *Sa*BirA-F123G binding to (F) *SabioO*, (G) *bioY* and (E) *yhfS-yhfT*. Concentrations of enzyme used in the binding reactions are shown in Figure 4.



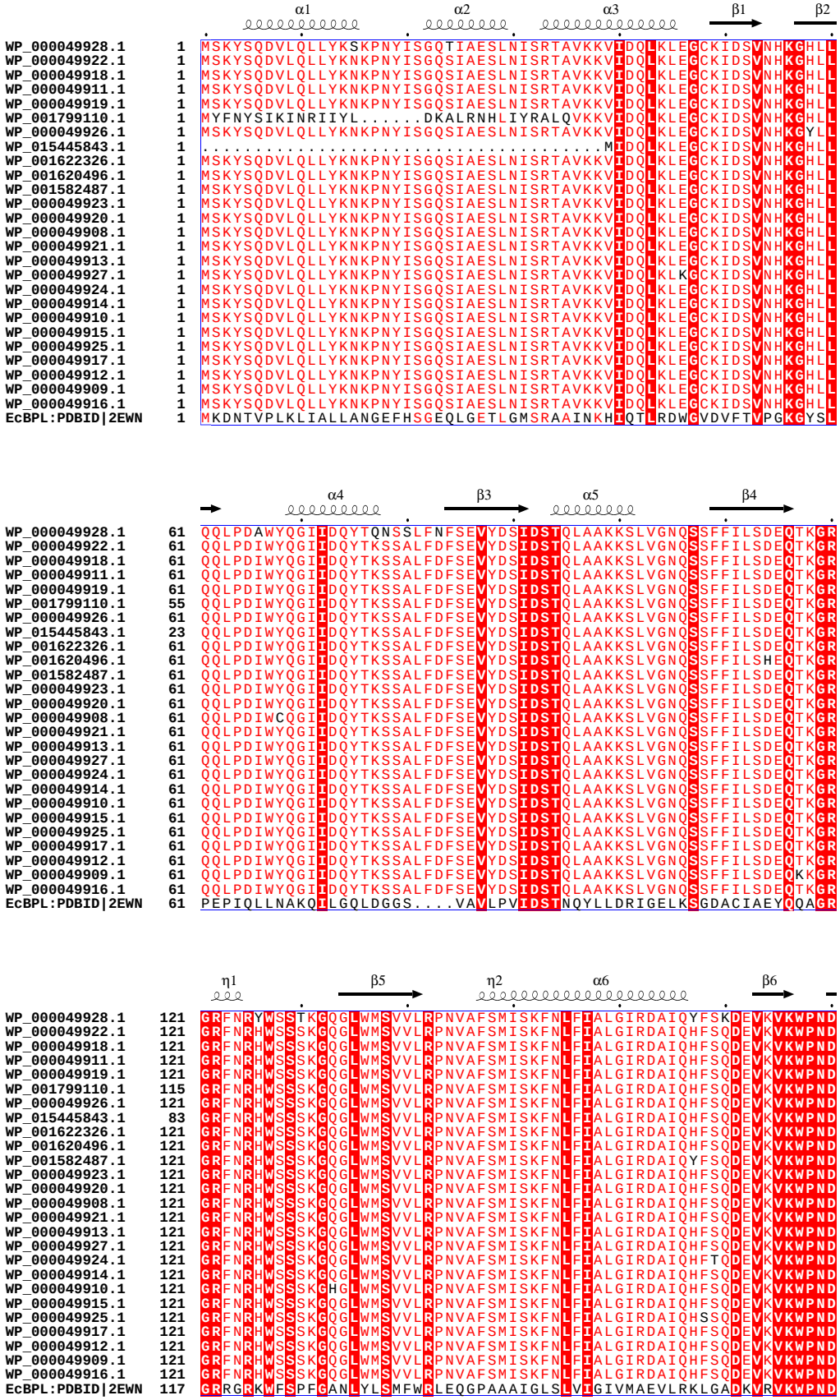
AATGTAAAC
5′-GCCTTACACAGACGATTATTAATTATAAAAGTTTACATTCGGAT-3′

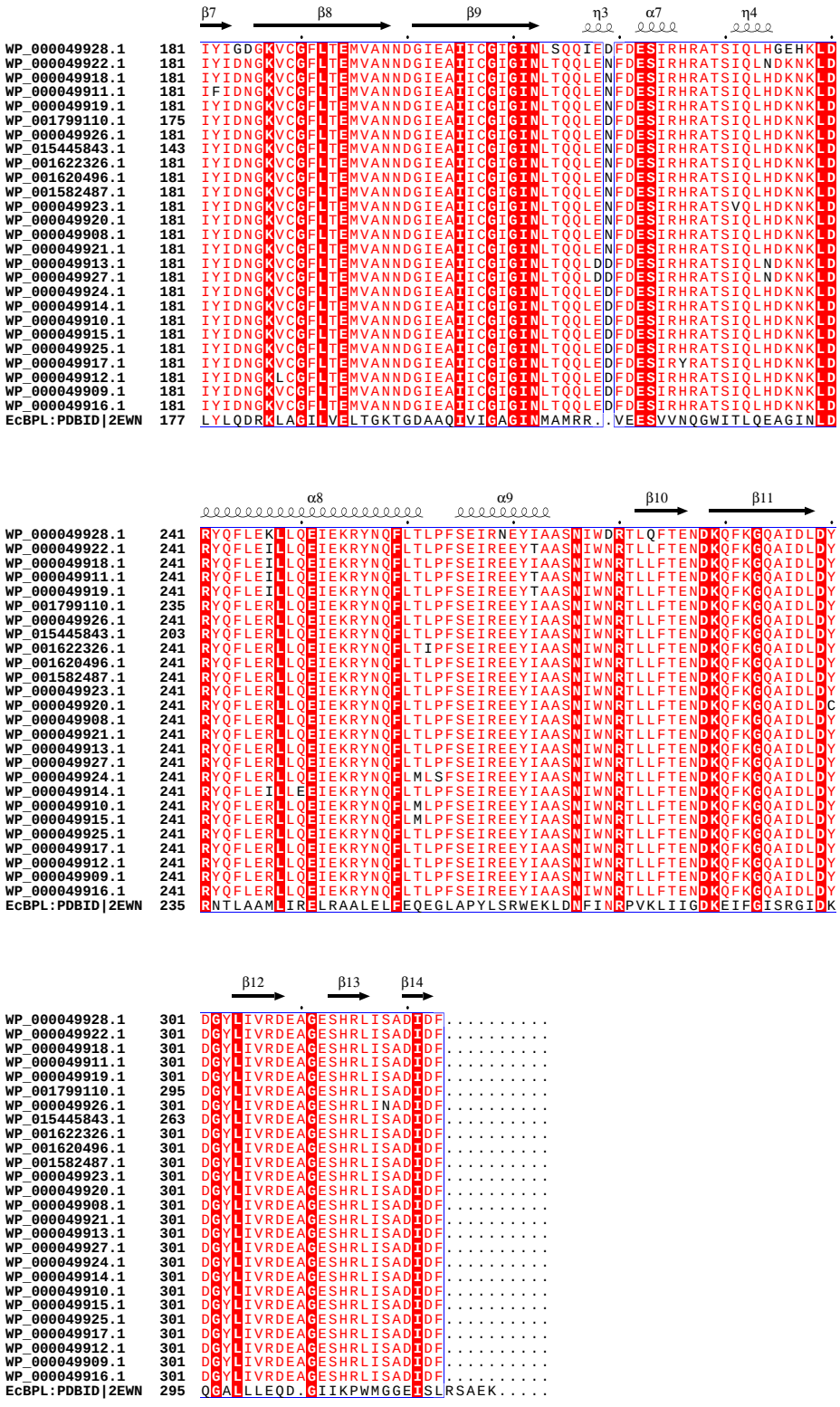


GTTTACATT
5′-GCCTTAAATGTAAACTTATTAATTATAAAAAGCAGACACCGGAT-3′

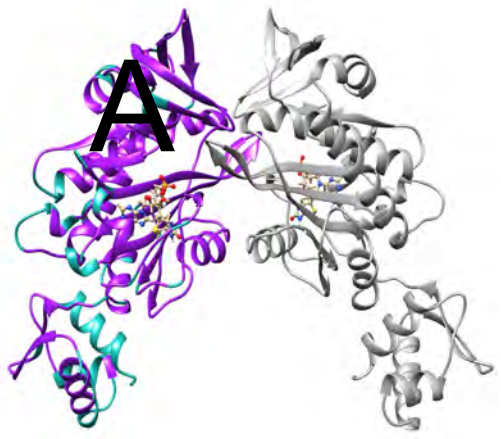
Supporting Figure S6. EMSA analysis of *Sa*BirA binding to a *SabioO* oligonucleotide containing a single BirA half site.

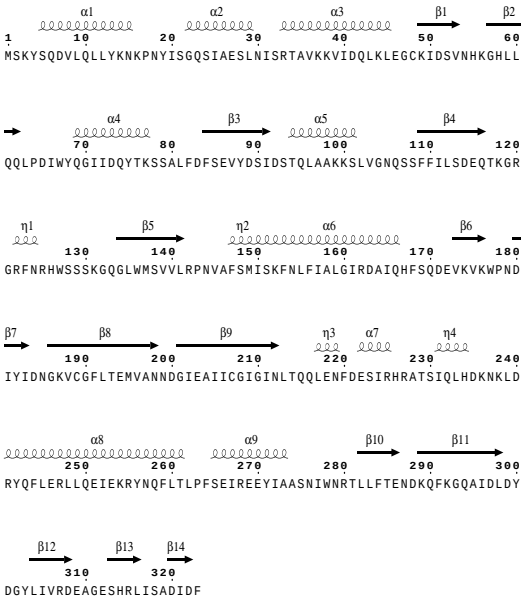
EMSA analysis was performed on double stranded *SabioO* oligonucleotides containing (top panel) the first half-site mutated and (lower panel) the second half-site mutated. The oligonucleotides are shown below the EMSA with the mutated sequence in red, the wild-type sequence is underlined and the half sites highlighted in yellow. For both oligonucleotides, no DNA binding was observed for *Sa*BirA (up to a concentration of 2 µM).


α1 α2 α3 β1 β2
WP_000049928.1 1 MSKYSQDVLQLLYKSKPNYISGQTIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049922.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049918.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049911.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049919.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_001799110.1 1 MYFNYSIKINRIIYL.....DKALRNHLIYRALQVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049926.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGYLL
WP_015445843.1 1 .....MIDQLKLEGCKIDSVNHKGHLL
WP_001622326.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_001620496.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_001582487.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049923.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049920.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049908.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049921.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049913.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049927.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLKGCKIDSVNHKGHLL
WP_000049924.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049914.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049910.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049915.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049925.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049917.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049912.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049909.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
WP_000049916.1 1 MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL
EcBPL:PDBID|2EWN 1 MKDNTVPLKLIALLANGEFHSGEQLGETLGMSRAAINKHIQTLRDWGVDVFTVPGKGYSL
α4 β3 α5 β4
WP_000049928.1 61 QQLPDAWYQGIIDQYTQNSSLFNFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049922.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049918.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049911.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049919.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_001799110.1 55 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049926.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_015445843.1 23 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_001622326.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_001620496.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSHEQTKGR
WP_001582487.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049923.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049920.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049908.1 61 QQLPDIWCQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049921.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049913.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049927.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049924.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049914.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049910.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049915.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049925.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049917.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049912.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
WP_000049909.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQKKGR
WP_000049916.1 61 QQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSLVGNQSSFFILSDEQTKGR
EcBPL:PDBID|2EWN 61 PEPIQLLNAKQILGQLDGGS....VAVLPVIDSTNQYLLDRIGELKSGDACIAEYQQAGR
η1 β5 η2 α6 β6
WP_000049928.1 121 GRFNRYWSSTKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQYFSKDEVKVKWPND
WP_000049922.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049918.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049911.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049919.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_001799110.1 115 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049926.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_015445843.1 83 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_001622326.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_001620496.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_001582487.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQYFSQDEVKVKWPND
WP_000049923.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049920.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049908.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049921.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049913.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049927.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049924.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFTQDEVKVKWPND
WP_000049914.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049910.1 121 GRFNRHWSSSKGHGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049915.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049925.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHSSQDEVKVKWPND
WP_000049917.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049912.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049909.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
WP_000049916.1 121 GRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAIQHFSQDEVKVKWPND
EcBPL:PDBID|2EWN 117 GRRGRKWFSPFGANLYLSMFWRLEQGPAAAIGLSLVIGIVMAEVLRKLGADKVRVKWPND



Supporting Figure S7. Alignment of 26 non-redundant *Sa*BirA sequences and the *E. coli* homologue. Invariant residues are boxed in red, whilst highly conserved residues are in red text. Black text indicates the presence of a polymorphism, or divergence from the *Ec*BirA sequence (bottom sequence in the alignment).





Position	Consensus	Strain	Polymorphism	Domain localisation	Structural Element	Possible effect
2	S	WP_001799110.1	Y	N-terminal DNA binding domain		
3	K	WP_001799110.1	F	N-terminal DNA binding domain		
4	Y	WP_001799110.1	N	N-terminal DNA binding domain		
5	S	WP_001799110.1	Y	N-terminal DNA binding domain	α-helix 1	
6	Q	WP_001799110.1	S	N-terminal DNA binding domain	α-helix 1	
7	D	WP_001799110.1	I	N-terminal DNA binding domain	α-helix 1	
8	V	WP_001799110.1	K	N-terminal DNA binding domain	α-helix 1	
9	L	WP_001799110.1	I	N-terminal DNA binding domain	α-helix 1	
10	Q	WP_001799110.1	N	N-terminal DNA binding domain	α-helix 1	
11	L	WP_001799110.1	R	N-terminal DNA binding domain	α-helix 1	
12	L	WP_001799110.1	I	N-terminal DNA binding domain	α-helix 1	
13	Y	WP_001799110.1	I	N-terminal DNA binding domain	α-helix 1	
14	K	WP_001799110.1	Y	N-terminal DNA binding domain	α-helix 1	
15	N	WP_000049928.1	S	N-terminal DNA binding domain	α-helix 1	
15	N	WP_001799110.1	L	N-terminal DNA binding domain	α-helix 1	
22	G	WP_001799110.1	D	N-terminal DNA binding domain	α-helix 2	
23	Q	WP_001799110.1	K	N-terminal DNA binding domain	α-helix 2	
24	S	WP_000049928.1	T	N-terminal DNA binding domain	α-helix 2	
24	S	WP_001799110.1	A	N-terminal DNA binding domain	α-helix 2	
25	I	WP_001799110.1	L	N-terminal DNA binding domain	α-helix 2	
26	A	WP_001799110.1	R	N-terminal DNA binding domain	α-helix 2	
27	E	WP_001799110.1	N	N-terminal DNA binding domain	α-helix 2	
28	S	WP_001799110.1	H	N-terminal DNA binding domain	α-helix 2	
30	N	WP_001799110.1	I	N-terminal DNA binding domain		
31	I	WP_001799110.1	Y	N-terminal DNA binding domain		
32	S	WP_001799110.1	R	N-terminal DNA binding domain		
33	R	WP_001799110.1	A	N-terminal DNA binding domain	α-helix 3	
34	T	WP_001799110.1	L	N-terminal DNA binding domain	α-helix 3	
35	A	WP_001799110.1	Q	N-terminal DNA binding domain	α-helix 3	
39	V	WP_015445843.1	M	N-terminal DNA binding domain	α-helix 3	
46	E	WP_000049927.1	K	N-terminal DNA binding domain		
58	H	WP_000049926.1	Y	N-terminal DNA binding domain	β-sheet 2	
66	I	WP_000049928.1	A	N-terminal DNA binding domain		

68	Y	WP_000049908.1	C	catalytic domain		
77	K	WP_000049928.1	Q	catalytic domain	α-helix 4	
78	S	WP_000049928.1	N	catalytic domain		
80	A	WP_000049928.1	S	catalytic domain		
83	D	WP_000049928.1	N	catalytic domain		
114	D	WP_001620496.1	H	catalytic domain	β-sheet 4	next residue (115) contacts biotinol-5′-AMP
117	T	WP_000049909.1	K	catalytic domain		next to the biotin-binding loop (residues 118-128)
126	H	WP_000049928.1	Y	catalytic domain		in biotin-binding loop, contacts biotinol-5′-AMP
130	S	WP_000049928.1	T	catalytic domain		near the biotin-binding loop (residues 118-128)
133	Q	WP_000049910.1	H	catalytic domain		
167	H	WP_000049928.1	Y	catalytic domain		
167	H	WP_001582487.1	Y	catalytic domain		
168	F	WP_000049925.1	S	catalytic domain		
169	S	WP_000049924.1	T	catalytic domain		
170	Q	WP_000049928.1	K	catalytic domain		
182	Y	WP_000049911.1	F	catalytic domain	β-sheet 7	
184	D	WP_000049928.1	G	catalytic domain		
185	N	WP_000049928.1	D	catalytic domain		
188	V	WP_000049912.1	L	catalytic domain	β-sheet 8	previous residue (187) contacts biotinol-5′-AMP
214	T	WP_000049928.1	S	catalytic domain		next residue (215) contacts biotinol-5′-AMP
217	L	WP_000049928.1	I	catalytic domain	3_{10}-helix 3	
218	E	WP_000049913.1	D	catalytic domain	3_{10}-helix 3	
218	D	WP_000049927.1	D	catalytic domain	3_{10}-helix 3	
219	D/N	All strains	D/N	catalytic domain	3_{10}-helix 3	next residue (220) contacts biotinol-5'-AMP
226	H	WP_000049917.1	Y	catalytic domain		next residues (227-229) contact biotinol-5'-AMP
231	I	WP_000049923.1	V	catalytic domain	3_{10}-helix 4	
234	H	WP_000049922.1	N	catalytic domain	3_{10}-helix 4	
234	H	WP_000049913.1	N	catalytic domain	3_{10}-helix 4	
234	H	WP_000049927.1	N	catalytic domain	3_{10}-helix 4	
235	D	WP_000049928.1	G	catalytic domain		
236	K	WP_000049928.1	E	catalytic domain		
237	N	WP_000049928.1	H	catalytic domain		
247	R	WP_000049928.1	K	catalytic domain	α-helix 8	
247	R	WP_000049922.1	I	catalytic domain	α-helix 8	
247	R	WP_000049918.1	I	catalytic domain	α-helix 8	
247	R	WP_000049911.1	I	catalytic domain	α-helix 8	
247	R	WP_000049919.1	I	catalytic domain	α-helix 8	
247	R	WP_000049914.1	I	catalytic domain	α-helix 8	
250	Q	WP_000049914.1	E	catalytic domain	α-helix 8	
261	T	WP_000049924.1	M	catalytic domain	α-helix 8	
261	T	WP_000049910.1	M	catalytic domain	α-helix 8	
261	T	WP_000049915.1	M	catalytic domain	α-helix 8	
262	L	WP_001622326.1	I	catalytic domain		
263	P	WP_000049924.1	S	catalytic domain		
269	E	WP_000049928.1	N	C-terminal cap	α-helix 9	
272	I	WP_000049922.1	T	C-terminal cap	α-helix 9	
272	I	WP_000049911.1	T	C-terminal cap	α-helix 9	
272	I	WP_000049919.1	T	C-terminal cap	α-helix 9	
279	N	WP_000049928.1	D	C-terminal cap		
283	L	WP_000049928.1	Q	C-terminal cap	β-sheet 10	
300	Y	WP_000049920.1	C	C-terminal cap		
318	S	WP_000049926.1	N	C-terminal cap		involved in the dimer interface

Supporting Figure S8. Polymorphisms present in the various *S. aureus* strains, and their potential impact due to their structural localization.

(A) Location of polymorphisms (cyan) is shown in the ribbon diagram of *Sa*BirA detected from the sequence alignment (Supporting Figure S7). (B) Sequence of the Mu50 strain *Sa*BirA with secondary structures from the crystal structure annotated. (C) Table of all the polymorphisms within the 26 non-redundant *Sa*BirA sequences identified from the multiple sequence alignment (Supporting Figure S7). The position of the mutations in the crystal structure is annotated.

Supporting Table S1: List of oligonucleotides employed in this study.

Oligo name	Sequence 5'-3'	Description
Cloning oligonucleotides		
B386	GACTAAAATGTTGAATCGCATTCTTATCCCTAAATCAATAAATAAATTAAATTTAGATATCATTGAGAATGC	Sequencing primer to sequence pTac-SaBirA in integration plasmid pIT4_TL152002
HK022-P1	GGAATCAATGCCTGAGTG	attp-HK022 PCR screening primer [4]
HK022-P2	ACTTAACGGCTGACATGG	attp-HK022 PCR screening primer [4]
HK022-P3	ACGAGTATCGAGATGGCA	attp-HK022 PCR screening primer [4]
HK022-P4	GGCATCAACAGCACATTC	attp-HK022 PCR screening primer [4]
Lambda P1	GGCATCACGGC AATATAC	attp-λ PCR screening primer [4]
Lambda P2	ACTTAACGGCTGACATGG	attp-λ PCR screening primer [4]
Lambda P3	GGGAATTAATTCTTGAAGACG	attp-λ PCR screening primer [4]
Lambda P4	TCTGGTCTGGTAG CAATG	attp-λ PCR screening primer [4]
B460_R119W_F	GGCCGTGGTCGCTGGGGTCGGAAATGG	Forward mutagenesis primer for EcBirA-R119W
B461_R119W_R	CCATTTCCGACCCCAGCGACCACGGCC	Reverse mutagenesis primer for EcBirA-R119W
B479	GACTCATCATGAAGGATAACACCGTGCCAC	Forward primer to clone EcBirA-R119W into integration plasmid pIT4_TL 152002
B320	ACTAGTGATAAGCTTAATGATGATGATGATGATGTCC	Reverse primer to clone EcBirA-R119W into integration plasmid pIT4_TL_152002
EMSA and native MS oligonucleotides		
DS-SaBioO oligo 1	CCTTAAATGTAAACTTTTATAATTAATAAGTTTACATTTAAG	Top strand oligo containing SabioO wildtype sequence
DS-SaBioO oligo 2	CCTTAAATGTAAACTTATTAATTATAAAAGTTTACATTTAAGG	Bottom strand oligo containing SabioO wildtype sequence
DS-SabioY oligo 1	AACTTATTGTAAACTTTTCATTTCTTAAAGTTTACAATGGTGCT	Top strand oligo containing SabioY wildtype sequence

DS-SabioY oligo 2	AGCACCATTGTAAACTTTAAGAAATGAAAAGTTTACAATAAGTT	Bottom strand oligo containing SabioY wildtype sequence
DS-yHFS-T oligo 1	TTATATAATGTTAACAAGATGTATTTTAAAGTTTACATTGAGTGA	Top strand oligo containing yHFS-T wildtype sequence
DS-yHFS-T oligo 2	TCACTCAATGTAAACTTTAAAATACATCTTGTTAACATTATATAA	Bottom strand oligo containing yHFS-T wildtype
DS-HS1m oligo 1	TTACACAGACGATTATTAATTATAAAAGTTTACATTCG	Top strand oligo containing mutated sequence of the first half-site of *SaBirA* recognition sequence for SabioO
DS-HS1m oligo 2	CGAATGTAAACTTTTATAATTAATAATCGTCTGTGTAA	Bottom strand oligo containing mutated sequence of the first half-site of SaBirA recognition sequence for SabioO
DS-HS2m oligo 1	TTAAATGTAAACTTATTAATTATAAAAAGCAGACACCG	Top strand oligo containing mutated sequence of the second half-site of SaBirA recognition sequence for SabioO
DS-HS2m oligo 2	CGGTGTCTGCTTTTTATAATTAATAAGTTTACATTTAA	Bottom strand oligo containing mutated sequence of the second half-site of SaBirA recognition sequence for SabioO
DS-HSm oligo 1	TTACACAGACGATTATTAATTATAAAAAGCAGACACCG	Top strand oligo containing mutated sequence of both half-site of SaBirA recognition sequence for SabioO
DS-HSm oligo 2	CGGTGTCTGCTTTTTATAATTAATAATCGTCTGTGTAA	Bottom strand oligo containing mutated sequence of both half-site of SaBirA recognition sequence for SabioO

Supporting Table S2. Plasmids employed in this study.

Plasmid	Description	Reference
pGEMT-EcBirA-H6	pGEMT plasmid containing EcBirA with 6x his-tag	[3]
pIT3_CLlacZ_Trim	Chromosomal integration plasmid (λ-attP, Cm^R, R6Kγ ori, lacZ)	[5]
pIT3-SH-152002	Chromosomal integration plasmid (HK022-attP, $Spec^R$, R6Kγ ori, ccdB, pUC ori)	Shearwin lab, Adelaide University
placZ_SH_Trim	Chromosomal integration plasmid (HK022-attP, $Spec^R$, R6Kγ ori, lacZ)	this study
pIT4_TL_152002	Chromosomal integration plasmid (λ-attP, Tc^R, R6Kγ ori, ccdB, pUC ori)	[4]
pGEMT-EcBirA-R119W-H6	pGEMT plasmid containing EcBirA-R119W-H6	this study
pGEMT-SaBirA F123G-H6	pGEMT plasmid containing SaBirA-F123G-H6	[3]
pET16b-EcBirA-R119W-H6	pET16b expression vector containing EcBirA-R119W-H6	this study
pET16b-SaBirA-F123G-H6	pET16b expression vector containing SaBirA-F123G-H6	[3]
pIT4_TL_SaBirA-H6	plac-UV5 fused with SaBirA-H6 sequence cloned into pIT4_TL_152002	[6]
pIT4_TL_SaBirA-F123G-H6	plac-UV5 fused with SaBirA-F123G-H6 sequence cloned into pIT4_TL_152002	this study
pIT4_TL_EcBirA-R119W-H6	plac-UV5 fused with EcBirA-R119W-H6 sequence cloned into pIT4_TL_152002	this study

Supporting Table S3: Bacterial strains employed in this study.

Strain name	Genotype	Description	Source
KP7600	KP7600 (F- lacIQ lacZdeltaM15 galK2 galT22 lambda- in (rrnD-rrnE)1	W3110 derivative	[7]
JD26186	*bioC::KanR*	*E. coli* KP7600 derivative with disrupted *bioC* gene	NBRP, Japan
JD26186 *birA::CAT*	*bioC::KanR birA::CAT*	JD28186 with CAT insertion in endogenous Ec*birA*	[6]
JD26186 *birA::CAT-SabioO-SaBirA*	*bioC::KanR birA::CAT (SabioO-lacZ)HK(placUV5-SaBirA)*	JD26186 birA::CAT with with *Sa*BioO-lacZ reporter integrated at HK022 att site and placUV5-*Sa*BirA integrated at lambda att site.	[6]
JD26186 *birA::CAT-SabioY-SaBirA*	*bioC::KanR birA::CAT (SabioY-lacZ)HK (placUV5-SaBirA)*	JD26186 birA::CAT with *Sa*BioY-lacZ reporter integrated at HK022 att site and placUV5-*Sa*BirA integrated at lambda att site.	[6]
JD26186 *birA::CAT-yhfST-SaBirA*	*bioC::KanR birA::CAT (yHFT-lacZ)HK (placUV5-SaBirA)*	JD26186 birA::CAT with yHFT-lacZ reporter integrated at HK022 att site and placUV5-*Sa*BirA integrated at lambda att site.	[6]
JD26186 *birA::CAT-SabioO-SaBirA-F123G*	*bioC::KanR birA::CAT (SabioO-lacZ)HK (placUV5-SaBirA-F123G)*	JD26186 birA::CAT with *Sa*BioO-lacZ reporter integrated at HK022 att site and placUV5-*Sa*BirA-F123G integrated at lambda att site.	this study
JD26186 *birA::CAT-SabioY-SaBirA-F123G*	*bioC::KanR birA::CAT (SabioY-lacZ)HK (placUV5-SaBirA-F123G)*	JD26186 birA::CAT with *Sa*BioY-lacZ reporter integrated at HK022 att site and placUV5-*Sa*BirA-F123G integrated at lambda att	this study
JD26186 *birA::CAT-yhfST-SaBirA-F123G*	*bioC::KanR birA::CAT (yHFT-lacZ)HK (placUV5-SaBirA-F123G)*	JD26186 birA::CAT with yHFT-lacZ reporter integrated at HK022 att site and placUV5-*Sa*BirA-F123G integrated at lambda att	this study

JD26186-*EcBioO-EcBirA*	*bioC::KanR (EcBioO-lacZ)HK*	JD26186 with *Ec*BioO-lacZ reporter chromosomally integrated at HK022 att site	[6]
JD26186 *birA::CAT-EcBioO-EcBirA-R119W*	*bioC::KanR birA::CAT (EcBioO-lacZ)HK (placUV5-EcBirA-R119W)*	JD26186 birA::CAT with *Ec*BioO-lacZ reporter and placUV5-*Ec*BirA-R119W integrated at lambda att	this study

Supporting Table S4: Oligomeric state validation by native nano-electrospray ionisation mass-spectroscopy (nESI-MS)

*Ec*BirA and *Sa*BirA were analyzed using nESI-MS under native conditions to determine the oligomeric state of the protein in the absence and presence of of biotin and MgATP. Holo-BirA was prepared by incubating the apo-purified proteins with 500 µM biotin and 1 mM MgATP, prior to buffer exchange for MS analysis. The measured molecular masses (Da) and the corresponding oligomeric states are outlined, along with the predicted masses (Da) for these states. All measured masses are within the acceptable error range of ~1% in 1 MDa [8].

BirA Sample	Measured MWs (Da)	Complex components	Calculated MWs (Da)
Apo-wild type *Ec*BirA	36203	Monomer	36192
Holo-wild type *Ec*BirA	36771 73559	Monomer, biotinyl-5′-AMP bound Dimer, biotinyl-5′-AMP bound	36765 73530
Apo-R119W *Ec*BirA	36216	Monomer	36222
Holo-R119W *Ec*BirA	36783	Monomer, biotinyl-5′-AMP bound	36795
Apo-wild type *Sa*BirA	37892	Monomer	37892
Holo-wild type *Sa*BirA	38470 76925	Monomer, biotinyl-5′-AMP bound Dimer, biotinyl-5′-AMP bound	38466 76931
Apo-F123G *Sa*BirA	37802	Monomer	37802
Holo-F123G *Sa*BirA	38381	Monomer, biotinyl-5′-AMP bound	38376

Supporting Table S5: Native nESI-MS measured masses for BirA-DNA complexes.

BirA and double-stranded oligonucleotide complexes were prepared prior to analysis by nESI-MS as described in materials and methods. The measured molecular weights (MW) for the species detected for each BirA-DNA complex were compared against the predicted sizes (calculated MW). Description of each species detected is outlined. The measured MWs were within the accepted mass differences of ~1% in 1 MDa [8].

BirA	Measured MW (Da)	Complex components	Calculated MW (Da)
Apo-*Ec*BirA (WT) + *Ec*bioO	30767	Free *Ec*bioO	30761
	36200	Apo-*Ec*BirA (WT) monomer	36192
	103138	2 Apo-*Ec*BirA (WT) subunits bound to *Ec*bioO	103145
Holo-*Ec*BirA (WT) + *Ec*bioO	30767	Free *Ec*bioO	30761
	72382	Apo-*Ec*BirA (WT) dimer	72384
	73522	Holo-*Ec*BirA (WT, biotinyl-5'-AMP bound) dimer	73530
	104272	2 Holo-*Ec*BirA (WT, biotinyl-5'-AMP bound) subunits bound to *Ec*bioO	104292
Apo-*Ec*BirA (R119W) + *Ec*bioO	30771	Free *Ec*bioO	30761
	36210	Apo-*Ec*BirA (R119W) monomer	36222
	103191	2 Apo-*Ec*BirA (R119W) subunits bound to *Ec*bioO	103205
Holo-*Ec*BirA (R119W) + *Ec*bioO	30768	Free *Ec*bioO	30761
	104348	2 Holo-*Ec*BirA (R119W, biotinyl-5'-AMP bound) subunits bound to *Ec*bioO	104351
Apo-*Sa*BirA (WT) + *Sa*bioY	27050	Free *Sa*bioY	27053
	102819	2 Apo-*Sa*BirA (WT) subunits bound to *Sa*bioY	102837
Holo-*Sa*BirA (WT) + *Sa*bioY	27050	Free *Sa*bioY	27053
	103989	2 Holo-*Sa*BirA (WT, biotinyl-5'-AMP bound) subunits bound to *Sa*bioY	103984
Apo-*Sa*BirA (F123G) + *Sa*bioY	27061	Free *Sa*bioY	27053
	102653	2 Apo-*Sa*BirA (F123G) subunits bound to *Sa*bioY	102657
Holo-*Sa*BirA (F123G) + *Sa*bioY	27063	Free *Sa*bioY	27053
	103806	2 Holo-*Sa*BirA (F123G, biotinyl-5'-AMP bound) subunits bound to *Sa*bioY	103804

Supporting Table S6: Native nESI-MS analysis for *Sa*BirA in complex with an oligonucleotide containing two mutated half sites.

*Sa*BirA was tested for binding specificity by incubating the repressor with a double-stranded oligonucleotide containing two mutated recognition sequences. The sequence of the oligonucleotide is shown in Supporting Figure S3. Measured molecular weights (MW) for the species detected were compared against the predicted sizes (calculated MW). Description of each species is outlined. The measured MWs were within the accepted mass differences of ~1% in 1 MDa [8].

BirA	Measured MW (Da)	Complex components	Calculated MW (Da)
Apo-*Sa*BirA (WT) + mutated oligonucleotide	23897	Free mutated oligonucleotide	23886
	37887	Apo-*Sa*BirA (WT) monomer	37892
Holo-*Sa*BirA (WT) + mutated oligonucleotide	23896	Free mutated oligonucleotide	23886
	37902	Apo-*Sa*BirA (WT) monomer	37892
	38474	Holo-*Sa*BirA (WT, biotinyl-5'-AMP bound) monomer	38466
Apo-*Sa*BirA (F123G) + mutated oligonucleotide	23897	Free mutated oligonucleotide	23886
	37817	Apo-*Sa*BirA (F123G) monomer	37802
Holo-*Sa*BirA (F123G) + mutated oligonucleotide	23889	Free mutated oligonucleotide	23886
	38370	Holo-*Sa*BirA (F123G, biotinyl-5'-AMP bound) monomer	38376

Supporting Table S7: Native nESI-MS measured masses for *Sa*BirA in complex with a *Sa*bioO oligonucleotide containing a single mutated half-site.

Complexes of *Sa*BirA and double-stranded *Sa*bioO oligonucleotides with the first half-site recognition sequence mutated were prepared as described in materials and methods. The sequence of the oligonucleotide is shown in Supporting Figure S4A. Measured molecular weights (MW) of the species detected were compared against the predicted sizes (calculated MW). Description of each species is outlined. The measured MWs were within the accepted mass differences of ~1% in 1 MDa [8].

BirA	Measured MW (Da)	Complex components	Calculated MW (Da)
Apo-*Sa*BirA (WT) + 1 half-site mutated *Sa*bioO	23901	Free oligonucleotide	23884
	37881	Apo-*Sa*BirA (WT) monomer	37892
	61773	1 Apo-*Sa*BirA (WT) subunit bound to 1 half-site mutated *Sa*bioO	61776
Holo-*Sa*BirA (WT) + 1 half-site mutated *Sa*bioO	23888	Free oligonucleotide	23884
	37899	Apo-*Sa*BirA (WT) monomer	37892
	38475	Holo-*Sa*BirA (WT, biotinyl-5'-AMP bound) monomer	38466
	76942	Holo-*Sa*BirA (WT, biotinyl-5'-AMP bound) dimer	76931
	100811	2 Holo-*Sa*BirA (WT, biotinyl-5'-AMP bound) bound to 1 half-site mutated *Sa*bioO	100815
Apo-*Sa*BirA (F123G) + 1 half-site mutated *Sa*bioO	23891	Free oligonucleotide	23884
	37812	Apo-*Sa*BirA (F123G) monomer	37802
	61687	1 Apo-*Sa*BirA (F123G) subunit bound to 1 half-site mutated *Sa*bioO	61686
Holo-*Sa*BirA (F123G) + 1 half-site mutated *Sa*bioO	23893	Free oligonucleotide	23884
	38385	Holo-*Sa*BirA (F123G, biotinyl-5'-AMP bound) monomer	38376
	62268	1 Holo-*Sa*BirA (F123G, biotinyl-5'-AMP bound) subunit bound to 1 half-site mutated *Sa*bioO	62260
	100638	2 Holo-*Sa*BirA (F123G, biotinyl-5'-AMP bound) subunits bound to 1 half-site mutated *Sa*bioO	100635

References:

1 Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D. & Matthews, B. W. Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution. *J Mol Biol* **357**, 509-523, doi:10.1016/j.jmb.2005.12.066 (2006).

2 Soares da Costa, T. P. *et al.* Selective inhibition of biotin protein ligase from Staphylococcus aureus. *J Biol Chem* **287**, 17823-17832, doi:10.1074/jbc.M112.356576 (2012).

3 Soares da Costa, T. P. *et al.* Dual roles of F123 in protein homodimerization and inhibitor binding to biotin protein ligase from Staphylococcus aureus. *Mol Microbiol* **91**, 110-120, doi:10.1111/mmi.12446 (2014).

4 St-Pierre, F. *et al.* One-step cloning and chromosomal integration of DNA. *ACS Synthet. Biol.* **2**, 537-541, doi:10.1021/s400021j (2013).

5 Cui, L., Murchland, I., Dodd, I. B. & Shearwin, K. E. Bacteriophage lambda repressor mediates the formation of a complex enhancer-like structure. *Transcription* **4**, 201-205 (2013).

6 Satiaputra, J. *et al.* Biotin-mediated growth and gene expression in *Staphylococcus aureus* is highly responsive to environmental biotin. *Appl Microbiol Biotechnol* **102**, 3793-3803, doi:10.1007/s00253-018-8866-z (2018).

7 Miki, T., Yamamoto, Y. & Matsuda, H. A novel, simple, high-throughput method for isolation of genome-wide transposon insertion mutants of Escherichia coli K-12. *Methods Mol Biol* **416**, 195-204, doi:10.1007/978-1-59745-321-9_13 (2008).

8 Benesch, J. L. P. & Ruotolo, B. T. Mass Spectrometry: an Approach Come-of-Age for Structural and Dynamic Biology. *Curr Opin Struct Biol* **21**, 641-649 (2011).

CHAPTER 4:

KINETIC AND STRUCTURAL CHARACTERISATION OF CLASS III BPLs AS ANTIFUNGAL TARGETS

CHAPTER OUTLINE

This manuscript-style chapter describes the identification and production of three different fungal class III BPLs. The enzymes were then characterised using low-resolution structural techniques and stability experiments, including MS experiments similar to those optimised with the *S. aureus* BPL in chapter 3. Enzymatic assays were also employed to measure the Michaelis constants for the substrates. *In vitro* inhibition of the fungal BPLs was also assessed using re-purposed BPL inhibitors designed to target the bacterial *S. aureus* BPL. Comparisons between the three different BPLs are discussed and related back to the sequence homology results. Overall, this chapter characterises three fungal BPLs and discusses the potential for targeting fungal BPLs for the development of novel antifungal therapeutics.

STATEMENT OF AUTHORSHIP

Statement of Authorship

Title of Paper	Characterisation of class III biotin protein ligases; novel antifungal targets for human and agricultural pathogens
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Louise M Sternicki
Contribution to the Paper	Assisted with project and experimental design. Produced recombinant proteins and completed all experiements unless otherwise stated. Data analysis for all experiments, writing and preparing the manuscript.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 6/3/2019.

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Stephanie Nguyen
Contribution to the Paper	Sub-cloned *B. cinerea* and *Z. tritici* BPLs for complementation and expression. Completed complementation assay, and developed expression and purification strategies for these enzymes. Produced these recombinant proteins. Optimised activity assay for these BPLs.
Signature	Date 06/03/19

Name of Co-Author	Nicole R Pendini
Contribution to the Paper	Developed expression and purification strategies for the *S. cerevisiae* BPL, including appropriate sub-cloning.
Signature	Date 18/6/2019

Name of Co-Author	Kamila Pacholarz
Contribution to the Paper	Advised on mass spectrometry experiments, assisted with mass spectrometry data collection and analysis.
Signature	Date 7/10/2019

Name of Co-Author	Perdita Barran
Contribution to the Paper	Advised on mass spectrometry experiments.
Signature	Date 24/7/19

Name of Co-Author	Tara L Pukala
Contribution to the Paper	Assisted with mass spectrometry data collection and analysis. Provided insights into the project and helped with manuscript preparation.
Signature	Date 14/6/19

Name of Co-Author	Grant W Booker
Contribution to the Paper	Assisted with project and experiment design, data analysis and manuscript preparation.
Signature	Date 18/6/2019

Name of Co-Author	Kate L Wegener
Contribution to the Paper	Advised on project and experimental design, data analysis and manuscript preparation.
Signature	Date 18-06-19

Name of Co-Author	Steven W Polyak
Contribution to the Paper	Conceived and supervised the project, assisted with manuscript preparation.
Signature	Date 14-06-2019

MANUSCRIPT

Characterisation of class III biotin protein ligases; novel antifungal targets for human and agricultural pathogens

Louise M Sternicki[1], Stephanie Nguyen[1], Nicole R Pendini[1], Perdita Barran[2], Kamila Pacholarz[2], Tara L Pukala[3], Grant W Booker[1], Kate L Wegener[1, 4] and Steven W Polyak[1, 4, 5] *

[1] School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

[2] Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, United Kingdom

[3] School of Physical Sciences, The University of Adelaide, South Australia 5005, Australia

[4] Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

[5] Current Address: School of Pharmacy and Medical Sciences, University of South Australia, South Australia 5001, Australia

*To whom correspondence should be addressed.

Abstract

Biotin protein ligase (BPL) is an essential enzyme for all forms of life, making it a promising target for novel anti-infective agents. Whilst, most bacteria and archaea have simple BPL structures (class I and II), eukaryotic BPLs (class III) have evolved a more complex structure that includes a large N-terminal extension, comprising just over half the enzyme, in addition to the biotinylation catalytic domain. The absence of an atomic resolution structure of a class III BPL hinders structural and functional analysis of this enzyme, however, a stable, folded domain has been predicted to lie within this extension, responsible for enhancing the catalytic activity of the BPL. Here, three different class III BPLs were identified and produced from the fungi *Saccharomyces cerevisiae, Botrytis cinerea* and *Zymoseptoria tritici*. Circular dichroism and ion mobility mass spectrometry revealed the overall tertiary and secondary structures of these enzymes were similar, corresponding with the high levels of sequence similarity. However, subtle structural differences between the enzymes were determined from the different stabilities and varied Michaelis constants for the interaction with substrates biotin, MgATP and different biotin domains. Interestingly, the BPLs had different preferences for fungal versus bacterial biotin domains, providing further evidence that class III BPLs have a 'substrate validation' activity for selecting only appropriate substrates for biotinylation. Finally, selective, potent inhibition of these three BPLs was demonstrated despite their sequence and structural homology. This highlights the potential for targeting BPL for novel, selective antifungal therapies against *B. cinerea, Z. tritici* and other fungal species.

Introduction

Fungal infections are a substantial contributor to microbial related deaths worldwide, resulting in over 1.5 million deaths annually (1,2). A variety of fungi can cause invasive, systemic infections, however, the four fungal genera *Cryptococcus, Candida, Aspergillus* and *Pneumocystis*, are the most common. Together they are responsible for more than 90% of deaths caused by invasive fungal infections (2). Fungal infections not only pose a morbidity and mortality threat to humans but are also a serious hazard for global crop and food production. *Botrytis cinerea*, for example, is a necrotroph that results in grey mould disease, infecting the leaves, stems, flowers and fruit of over 1000 plant species worldwide including vital crops, berries, tomatoes and ornamental flowers (3-5). This important pathogen is responsible for pre- and post-harvest decay with yield losses of 15-40% through post-harvest spoilage (5). The economic damage is estimated at €1 billion annually (~$1.56 billion AUD at time of writing) (3). The fungal pathogen *Zymoseptoria tritici* causes the wheat disease Septoria tritici blotch, which is the most damaging disease of wheat in Europe (6,7). It is a global threat to many wheat production areas, and can be responsible for yield losses of 20% to 50% in untreated fields and even up to 10% in treated fields (7,8). Approximately 70% of the cereal crop fungicide usage in the European Union is utilised to control *Z. tritici*, costing $1.2 billion USD annually (~$1.7 billion AUD at time of writing) (7-9). The growing resistance of fungi to treatments and antifungal agents in both the healthcare and agricultural settings will only add to the detrimental effects of these human and plant pathogens (2,4-16). Therefore, new antifungal agents with novel mechanisms of action are required to combat these fungal pathogens in each setting.

Pharmacological inhibition of the utilization of biotin represents a promising new approach for the discovery of novel antifungal products. Biotin (Vitamin H or B7) is an essential vitamin for all forms of life. It serves as a cofactor for biotin-dependent enzymes, allowing these metabolic enzymes to carryout carboxyl transfer reactions (reviewed (17)). All living organisms possess between 1 and 6 biotin-dependent enzymes, most of which are involved in important metabolic roles for viability and/or virulence. Therefore, all organisms must also contain a biotin protein ligase (BPL) that is responsible for the covalent attachment of biotin onto these biotin-dependent enzymes (18). Protein biotinylation is catalysed by BPL using a two-step reaction mechanism to attach biotin onto a single conserved lysine residue located in the biotin carboxyl carrier domain (BCCD) of the biotin-dependent enzymes. The first partial reaction involves the binding of substrates biotin and ATP in adjacent pockets of the BPL to allow the formation of the biotinyl-5′-AMP reaction intermediate (19). BPL then forms protein-

protein interactions with the biotin carboxyl carrier domain (BCCD) of the target biotin-dependent enzyme such that the target lysine is precisely positioned in the BPL active site to facilitate the transfer of biotin onto the biotin-dependent enzyme (20). Structural and biochemical studies have revealed that the BPLs from *Staphylococcus aureus* (21), *Mycobacterium tuberculosis* (22,23), *Escherichia coli* (24) and *Homo sapiens* (25) possess an ordered binding mechanism whereby the biotin binds prior to ATP in the first partial reaction. Here, biotin binding is required to initiate conformational changes to disordered loops in the substrate-binding site. The correctly positioned residues then form the pocket necessary for ATP binding to occur (24).

Bacterial and archaeal BPLs have been extensively genetically, structurally and biochemically characterised, allowing thorough knowledge of the function of these BPLs. Structures of the BPLs from *E. coli* (24,26-28), *S. aureus* (21,29,30), *M. tuberculosis* (22,31-33), *Pyrococcus horikoshii* (20,34) and *Aquifex aeolicus* (35) have been solved. However, these enzymes represent only 2 of the 3 structural classes of BPLs: namely the class I and II enzymes. Class I enzymes are the smallest BPLs containing only the catalytic domain and the C-terminal cap structures, which are the minimal requirements for biotinylation. Class II enzymes contain an additional N-terminal DNA binding domain that allows these BPLs to transcriptionally regulate biotin synthesis and import (reviewed (36)). BPLs from these two classes are found in archaea and bacteria. Eukaryotic BPLs form the third structural class, the class III BPLs. The unique feature of this class is an elongated N-terminal extension, of roughly equivalent size to the catalytic domain and C-terminal cap, that does not share homology with the class II N-terminal DNA binding domain and is predicted not to bind DNA. There is currently no atomic resolution structure of a class III BPL nor this unique structural component, and the precise function of this extension has not been assigned. Limited structural and functional data exist for eukaryotic class III BPLs, with previous literature revealing the N-terminal extension is vital for complete biotinylation activity (37,38), and is hypothesised to enhance activity by enabling enduring interactions with the substrate BCCD for biotinyl-transfer to occur (25,38). These interactions are also suggested to provide specificity, allowing biotinylation of only appropriate substrates via a mechanism of ‘substrate verification’ (39).

BPLs are essential proteins for the viability of multiple organisms. Several studies demonstrate the importance of BPL for bacterial viability (32,40,41), and as such the BPLs from *S. aureus* and *M. tuberculosis* have been targeted for novel antibiotic design (21,30,32,33,42-44). BPL is also hypothesised to be an essential enzyme for fungi. The

necessity of BPL for *Saccharomyces cerevisiae* viability has been demonstrated through single allele knockouts (45) and the *Saccharomyces* Genome Deletion Project (46). When one allele of BPL was knocked out in *S. cerevisiae,* only half of the spores were viable post-sporulation with none of these spores being *bpl1* negative (45). Likewise, when the open reading frame encoding BPL was deleted in *S. cerevisiae* as part of a whole genome screen for genes associated with viability, the organism was no longer able to survive (46). BPL has also been demonstrated as an essential gene for the viability of the fungal organisms *Candida albicans* and *Schizosaccharomyces pombe* (47,48). Acetyl-CoA carboxylase 1 (ACC1), one of the biotin-dependent enzymes in *S. cerevisiae* that requires biotinylation for activity, is also essential for viability as it is necessary for the synthesis of lipids required for biosynthesis and maintenance of the cell membrane (49). Thus, inhibition of BPL also inhibits ACC1, preventing the activities of two essential gene products. It is noteworthy here, that disruption of the cell membrane and wall is a common mechanism of antifungals licenced for clinical and agricultural use (12,14). The other biotin-dependent enzymes are involved in energy metabolism pathways including the utilisation of urea as a nitrogen source (50,51), with energy production pathways already being targeted by current agricultural fungicides (12,52). Hence, BPL is an attractive target for the development of novel antifungal therapeutics, as its disruption would influence important fungal growth and virulence pathways (51,53). Furthermore, BPL is a promising target as its inhibition would inhibit multiple essential pathways at the same time, rather than just one pathway as for current antifungals. Whilst the requirement for a functional BPL has been exploited in the design of antibacterials, BPL has not been targeted previously for antifungal therapeutics.

Here we have characterised class III fungal BPLs from the agricultural pathogens *B. cinerea* and *Z. triciti*, alongside the BPL from the model yeast *S. cerevisiae*. Circular dichroism (CD) and ion mobility-mass spectrometry (IM-MS) experiments on purified recombinant protein revealed that all three class III BPLs share similar global structural folds. However, collision-induced unfolding-mass spectrometry (CIU-MS) and thermal denaturation studies revealed variable in-solution and gas phase stabilities, implying subtle structural differences between the enzymes. Similarly, activity assays revealed differences in Michaelis constants for the substrates biotin, ATP and various biotin domains between the three enzymes. Biochemical studies also identified chemical analogues of the BPL reaction intermediate as potent, selective small molecule inhibitors of these BPLs. Characterising these BPLs represents the initial stage of antifungal drug design efforts to develop desperately required novel therapeutics and fungicides for the clinic and agriculture.

Materials and Methods

Identification of *Bc*BPL and *Zt*BPL coding sequences

The *S. cerevisiae* BPL sequence was used to search against the genomes of *B. cinerea* and *Z. triciti* to identify their BPL sequences using BLAST (Basic Local Alignment Search Tool).

Sequence alignment of class III BPLs

Clustal Omega was utilised to align the sequences of the human, *S. cerevisiae, B. cinerea, Z. tritici*, *S. aureus*, *E. coli* and *M. tuberculosis* BPLs, with further manual inspection using BioEdit (54). The Ident and Sim tool from the online Sequence Manipulation Suite (55) was used to calculate the percentage identity and similarity between the sequences, using the default residue grouping to determine similarity. The active site residues that form interactions with the reaction intermediate, biotinyl-5′-AMP, were identified from the structures of the biotinyl-5′-AMP-bound *M. tuberculosis* (PDB: 4OP0) (22) and *S. aureus* BPLs (PDB:3RIR) (29) using LigPlot+ (56,57).

Complementation assay of class III BPLs

The identified sequences for *Bc*BPL and *Zt*BPL were commercially cloned into the pESC-Ura vector (Genescript). The coding sequences were excised from this vector by *Nco*I/*Hind*III restriction digests, before ligation into the similarly digested pAra13 expression vector. The resulting pAra13-*Bc*BPL and pAra13-*Zt*BPL vectors were transformed into *E. coli* BM4062. The pAra13 plasmid containing *Sc*BPL was previously transformed into *E. coli* BM4062 and stored as a glycerol stock (37). A strain containing the class III *Candida albicans* BPL was included as a positive control, having previously been characterised by this complementation assay (58). The complementation assay was completed by streaking a single colony on an LB agar plate supplemented with 200 µg/mL ampicillin to be grown at 42 °C, then streaking the same colony on a replicate plate to be grown at 30 °C. Plates were incubated overnight.

Cloning and transformation of pVT(*Sc*BPL-H_6) in *Saccharomyces cerevisiae* w303

pET(*Sc*BPL-H_6) (37) was digested with *Xba*I and *Xho*I endonucleases. The resulting 2 kb fragment was ligated into *Xba*I-*Xho*I digested pVT100 vector (59). pVT(*Sc*BPL-H_6) was transformed into competent *S. cerevisiae* w303 cells as previously described (60).

Recombinant production of *Sc*BPL-H_6

*Sc*BPL expression and purification was performed similarly to that described in (58). A 10 mL starter culture of w303 *S. cerevisiae* harbouring the pVT100u expression vector for *Sc*BPL-H_6 was grown in yeast minimal media (8 g/L Yeast Nitrogen Base without amino acids and ammonium sulphate (Difco), 11 g/L Cassamino Acids (Difco), 1.48 mM adenine, 0.01% tryptophan, 0.01% leucine, 2% glucose, 41 µM *d*-biotin) at 30 °C for 24 to 48 hours with agitation at 170 rpm. A 1:100 dilution of this starter culture was used to inoculate a 100 mL sub-culture of yeast minimal media that was grown at 30 °C for 24 hours with agitation at 170 rpm. Cells were pelleted by centrifugation at 3300 x *g* for 5 min before re-suspension in 10 mL yeast minimal media for the inoculation of 500 mL of fresh yeast minimal media, which was then grown at 30 °C for 48 hours with agitation at 170 rpm. Cells were harvested by centrifugation at 1700 x *g* at 4 °C for 10 minutes before freezing the pellets at -80 °C. The cell pellets were washed with 1x PBS and re-harvested by repeating the centrifugation before freezing the pellets at -80 °C again. Cells were then washed in wash buffer 1 (10 mM Tris pH 7.5, 0.65 M sorbitol, 0.1 mM EDTA, 0.1 mM DTT) three times, before being resuspended in wash buffer 2 (10 mM Tris pH 7.5, 0.65 M sorbitol, 0.1 mM EDTA, 1 mM DTT) supplemented with 1000 units of lyticase. This cell suspension was incubated for 1.5 hours at 30 °C with agitation at 150 rpm, before re-harvesting by centrifugation at 1900 x *g* for 15 minutes.

Cell lysates were prepared by chemical lysis using 5 mL of Yeast Buster ® (Novagen) per gram of cell pellet together with 1 mM Tris(hydroxypropyl)phosphine, 1 mM PMSF and cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche) at room temperature with gentle agitation for 1 hour. Lysate was then clarified by centrifugation at 1900 x *g* for 15 minutes and 0.45 µm filtration. Clarified lysate was applied to a 5 mL HiPrep™ IMAC Fast Flow column (GE Healthcare) pre-charged with nickel sulphate and equilibrated with NiNTA wash buffer (300 mM NaCl, 50 mM NaH_2PO_4, 20 mM imidazole) for purification of the hexa-histidine tagged *Sc*BPL. Following application of the lysate, the column was washed with NiNTA Wash Buffer, followed by 50 mM imidazole using 10% NiNTA Elution Buffer (300 mM NaCl, 50 mM NaH_2PO_4, 500 mM imidazole) in NiNTA Wash Buffer. Finally, *Sc*BPL-H_6 was eluted with 200 mM imidazole, using 40% NiNTA Elution Buffer in NiNTA Wash Buffer. IMAC-purified *Sc*BPL was buffer exchanged into BPL Storage Buffer (50 mM Tris pH 8.0, 5% glycerol, 1 mM DTT, 0.5 mM EDTA pH 8.0) via dialysis overnight at 4°C.

*Sc*BPL was resolved from residual contaminating proteins by anion exchange chromatography. The dialysed *Sc*BPL was loaded onto an 8 mL ENRICH Q-Sepharose

column (Bio-Rad) that was pre-equilibrated with BPL Storage Buffer. Following column washing with BPL Storage Buffer, *Sc*BPL was eluted with a 0-20% gradient of BPL Storage Buffer with 1 M NaCl. The fractions containing *Sc*BPL were pooled and buffer exchanged via dialysis overnight at 4 °C into BPL Storage Buffer to remove excess NaCl. *Sc*BPL was concentrated using Vivaspin® 20 30,000 MWCO PES spin concentrators (Sartorius) and stored at -80 °C. SDS-PAGE was utilised to assess the purity of *Sc*BPL.

Apo preparations of *Sc*BPL-H_6 were prepared by treatment of the clarified cell lysate with streptaviding agarose, then a biotin accepting substrate (namely the biotin domain from the *S. cerevisiae* pyruvate carboxylase isoform 1) prior to purification, as outlined in (61).

Recombinant production of *Bc*BPL and *Zt*BPL

The sequences of *Bc*BPL and *Zt*BPL cloned into the pESC-Ura vector were purchased from Genescript. The *Nco*I/*Hind*III digested 2.1 kb *Bc*BPL and 2.2 kb *Zt*BPL fragments were sub-cloned into *Nco*I/*Hind*III digested pET-H_6-TEV vector for expression in *E. coli.* The resulting positive clones were transformed into *E. coli* BL21 (λDE3)-RIPL cells for expression in LB broth supplemented with 100 µg/mL ampicillin and 50 µg/mL chloramphenicol. An overnight culture was diluted 1:20 into fresh LB and grown at 30 °C with agitation until an OD_{600} of 0.6-0.8 was reached. Protein expression was induced with 0.2 mM IPTG and incubated at 16 °C with agitation overnight. Cells were harvested by centrifugation at 4500 x *g* for 10 minutes at 4 °C. The cell pellet was washed with PBS and the cells re-harvested by repeating the centrifugation.

Cells were resuspended in IMAC Buffer A (20 mM Tris, 500 mM NaCl, 1 mM BME, 10 mM Imidazole) then lysed with a minimum of 5 passages through a M110L homogeniser (Microfluidics). The lysate was clarified with 2 centrifugation steps of 10 minutes each at 15,000 x *g* and 4 °C, before filtering through a 0.45 µM filter. The clarified, soluble lysate was loaded onto a 5 mL Nickel-NTA chelating column (Bio-Scale™ Mini Profinity™ IMAC Cartridge, Bio-Rad Laboratories) that had been pre-equilibrated with IMAC Buffer A. Non-specifically bound proteins were removed by washing with IMAC Buffer A and then 50 mM imidazole using 20% IMAC Buffer B (20 mM Tris, 500 mM NaCl, 1 mM BME, 250 mM Imidazole) in IMAC Buffer A. BPLs were eluted with 100% IMAC Buffer B and fractions containing *Bc*BPL or *Zt*BPL were pooled for size exclusion chromatography to further purify and buffer exchange the enzymes. A 300 mL gel filtration column (HiPrepTM 26/60 Sephacryl® S-300 High Resolution (GE Healthcare)) was pre-equilibrated with 330 mL of BPL

Storage Buffer (50 mM Tris, 5% glycerol, 1 mM EDTA, 1 mM DTT) before applying *Bc*BPL or *Zt*BPL. Protein was eluted with 1.2 column volumes of BPL Storage Buffer. Fractions containing BPL were pooled, concentrated using Vivaspin® 20 30,000 MWCO PES spin concentrators (Sartorius) and stored at -80 °C. SDS-PAGE was utilised to analyse the purity of the BPLs.

Denatured Mass Spectrometry for accurate mass measurements of purified BPLs

Protein (buffer exchanged into 200 mM ammonium acetate using Amicon® Ultra 0.5 mL Centrifugal Filters (Merck)) was diluted to 2 µM in 50% acetonitrile + 0.1% formic acid. Denatured protein mass measurements were completed using an Agilent 6230 Time of Flight Mass Spectrometer (ToF MS) instrument (Agilent) with a dual Agilent Jet Stream electrospray ionisation (AJS ESI) source coupled to an Agilent 1260 Infinity Liquid Chromatography (LC) system. Protein was directly injected (5 µL) without chromatographic separation, and electrosprayed with 50% aqueous acetonitrile/0.1% formic acid at a flow rate of 0.2 mL/min. Electrospray ionisation-mass spectrometry (ESI-MS) conditions included: positive ion mode; capillary voltage, 3500 V; nozzle voltage, 1000 V; fragmentor, 360 V; gas, 8 L/min; gas temperature, 250°C; sheath gas, 12 L/min; and sheath gas temperature, 275°C. Spectra were deconvoluted for mass determination using BioConfirm software (Agilent).

Native ion mobility-mass spectrometry (IM-MS) of class III BPLs

BPLs often co-purify as a mixture of biotin-bound, reaction intermediate-bound and apo protein. In order to have a homogenous sample of each BPL, they were incubated with 0.5 mM biotin, 1 mM $MgCl_2$ and 1 mM ATP on ice for at least 1 hour to promote the formation of reaction intermediate-bound BPL. Proteins were then buffer exchanged into 100 mM ammonium acetate pH 6.9 using three sequential Micro Bio-Spin™ P-6 Gel Columns (Bio-Rad). Proteins were then quantified using a NanoDrop (ThermoFisher Scientific), with the required molecular weight and extinction coefficients calculated using the ProtParam Tool on the ExPASy server (62). Proteins were diluted to 5 or 10 µM in 100 mM ammonium acetate for nESI-MS. β-lactoglobulin, avidin, albumin, concanavalin A, and alcohol dehydrogenase were also buffer exchanged and quantified as above, and diluted to 10 µM in 100 mM ammonium acetate. These proteins were used as collision cross section (CCS) calibrants (63,64).

MS measurements were completed using a G2-S Synapt High Definition Mass Spectrometer (HDMS) (Waters Corporation). Samples were introduced in the positive ion mode from glass

capillaries (prepared in house) containing an inserted platinum wire (Goodfellow, diameter 0.125 mm) via a nano-electrospray ionisation (nESI) source. Conditions were optimised to maintain non-covalent interactions and native protein structure, and included capillary voltage, 1.0 to 1.9 kV; cone voltage, 50 V; source temperature, 40 °C; trap collision energy, 4 V; transfer collision energy, 2 V; trap gas, 2 mL/min; and backing pressure, 3.67 mbar. Specific ion mobility (IM) parameters included IM cell wave height, 35 V; IM cell wave velocity, 400 m/s; transfer t-wave height, 0 V; and transfer t-wave velocity, 191 m/s.

MassLynx V4.1 and DriftScope^{TM} 2.8 (Waters Corporation) were utilised to extract the ion drift times for CCS determination utilising previously described calculations (64).

Collision induced unfolding-mass spectrometry (CIU-MS)

Samples were prepared and introduced into the Synapt G2-S mass spectrometer as described above for native IM-MS. Data were collected in the MSMS mode, selecting the *m/z* of the charge state of interest. ORIGAMI^{MS} software (65) was utilised for data collection. Trap collision energies began at 4 V and were raised in 2 V increments to 200 V. At each voltage 5 scans were collected, with each scan 5 seconds in length. Data were analysed via ORGIAMI^{ANALYSE} (65).

Solution thermal denaturation to determine melting temperature

Thermal denaturation assays were completed according to Boivin *et al.* (66) and Reinhard *et al.* (67). Briefly, 10 µM of enzyme was incubated with 50 µM biotin, 100 µM $MgCl_2$ and 100 µM ATP on ice for 30 minutes. Protein was diluted 10-fold to a final concentration of 1 µM in 100 mM ammonium acetate and 5x SYPRO® Orange protein gel stain (ThermoFisher Scientific) in a total volume of 25 µL in the wells of an Applied Biosystems MicroAmp^{TM} Fast Optical 96-well reaction plate (ThermoFisher Scientific). Each enzyme was plated in triplicate, together with an enzyme only control for each protein and a dye only control. Plates were centrifuged at low speed (approximately 1000 x *g*) for 1 minute before sealing with an Applied Biosystems MicroAmp^{TM} Optical Adhesive Film plate seal (ThermoFisher Scientific). A Step One Plus qPCR instrument (Applied Bioscience) was used to perform the melting curve, measuring fluorescence using the ROX reporter channel every 0.3 °C, from 4 °C to 90 °C. Data were exported into Microsoft Excel before using GraphPad Prism to fit a Boltzman-Sigmoidal equation to the melting curve. This fit determined the melting point (T_M), which is the inflection point (V_{50}) of the melting curve. Each replicate was independently fit, and then the triplicate T_M values were averaged.

Circular Dichroism (CD) of class III BPLs

Proteins were buffer exchanged using Amicon Ultra-0.5 MWCO 10,000 centrifugal filter units (Merck) into 10 mM ammonium acetate pH 6.9 with 50 µM biotin, 100 µM $MgCl_2$ and 100 µM ATP. Concentrations were determined using a NanoDrop (ThermoFisher Scientific) (as for native IM-MS) and proteins diluted to 2.5 µM. Far UV CD measurements were completed on a Jasco CD J-815 spectropolarimeter at room temperature. Spectra were collected from 195 nm to 260 nm with 0.2 nm increments utilising a 10 mm path length cell. Results are the average of 5 scans corrected with buffer blanks and smoothed using Savitzky-Golay convolution. Spectra were exported from Jasco Spectral Analysis software for analysis with Microsoft Excel and GraphPad Prism.

Recombinant production of Apo-GST-*Sc*PC104

The coding region for the C-terminal 104 amino acids of the *S. cerevisiae* pyruvate carboxylase-1 (68) was subcloned into pGEX-4T-2 for expression as a GST fusion protein (GST-*Sc*PC104). This vector was transformed into the *birA1* conditional mutant strain *E.coli* BM4062 (69). An overnight culture of a successfully transformed colony was grown in LB with 200 µg/ml ampicillin at 30 °C with agitation. The overnight culture was utilised to inoculate fresh LB supplemented with 200 µg/ml ampicillin with a 1/50 dilution. Subcultures were grown at 30 °C with agitation until they reached mid-log phase (OD_{600} of 0.6-0.8), where they were then incubated at 42 °C for 30 minutes to inactivate the endogenous temperature-senstive mutant variant of *E. coli* BPL. Protein expression was then induced with 300 µg/ml IPTG for 3 hours at 42 °C. Cells were harvested by centrifugation at 4500 x *g* for 10 minutes at 4 °C. The cell pellet was washed with PBS and re-harvested by repeating the centrifugation.

The cell pellet was re-suspended in TBS pH 8.5, then lysed by passing the suspension a minimum of 5 times through a M110L homogeniser (Microfluidics). The cell lysate was clarified with two 10 minute centrifugation steps, each at 15,000 x *g* and 4°C, before filtering the soluble lysate through a 0.45 µM filter. Apo-GST-*Sc*PC104 was purified via a XK 26/20 column (Pharmacia, 20 cm x 2.6 xm) packed with 50 mL glutathione agarose (Sigma-Aldrich) and equilibrated with TBS pH 8.5 prior to loading the clarified lysate. The column was washed with TBS pH 8.5, before elution of GST-*Sc*PC104 with 10 mM reduced glutathione in TBS pH 8.5. Eluted Apo-GST-*Sc*PC104 was dialysed against 4 L of TBS pH 8.5 overnight at 4 °C to remove excess reduced glutathione, prior to quantification by Bradford assay. SDS-PAGE analysis was used to determine GST-*Sc*PC104 purity, before protein concentration using Vivaspin® 20 30,000 MWCO PES spin concentrators (Sartorius) and storage at -80 °C.

Recombinant production of Apo-GST-*Ca*PC115, GST-*Bc*PC103 and GST-*Sa*PC90

The biotin domain from the 115 C-terminal residues of *C. albicans* pyruvate carboxylase, GST-*Ca*PC115, was produced as above for Apo-GST-*Sc*PC104 (58). The *B. cinerea* pyruvate carboxylase biotin domain, contained in the C-terminal 103 residues, was purchased cloned into pGEX-4T-2 (Genescript). GST-*Bc*PC103 was transformed into the *E. coli* BL21 (DE3) strain for production as above. Similarly, the biotin domain from the 90 C-terminal residues of the *S. aureus* pyruvate carboxylase, GST-*Sa*PC90, was produced as above but also expressed in the *E. coli* BL21 (DE3) strain. Therefore, for expression of GST-*Bc*PC103 and GST-*Sa*PC90 the cultures were not required to be cultured at 42 °C, but rather induced at mid-log phase and grown for a further 3 hours at 30 °C.

Biotinylation blot to measure amount of biotinylated biotin domain substrate

A western blot probed with Streptavidin conjugated to the fluorophore Alexa488 was utilised to measure the endogenous biotinylation levels of the purified biotin domain-GST fusion proteins that occurred during expression. The blot was performed according to (61,70). The reactions completed prior to the blot contained 15 µM of the biotin domain substrate, except for GST-*Bc*PC103 were 13.8 µM was used, and 3 µM of BPL for the biotinylated controls.

Biotinylation activity assays to measure Michaelis constants for substrates

Activity assays to measure the Michaelis constants of the different BPLs for substrates were performed according to the time resolved fluorescence method outlined previously (30). Briefly, assay reactions contained 50 mM Tris-HCl pH 8.0, 3 mM ATP, 5.5 mM $MgCl_2$, 15 µM biotin, 0.1 µM DTT and 15 µM of appropriate biotin domain substrate. The reaction was initiated by the addition of BPL to a final concentration of 4.5 nM for *Sc*BPL, or 35 nM for *Bc*BPL and *Zt*BPL. The reaction, carried out at 37 °C, was terminated after 15 minutes for *Sc*BPL and *Bc*BPL, and 25 minutes for *Zt*BPL by the addition of 80 µL of stop buffer (55 mM EDTA, TBS). Aliquots of 20 µL were plated in triplicate in the wells of a white Lumitrac-600 96-well plate (Greiner) that was pre-coated with a polyclonal anti-GST antibody (Sigma-Aldrich). To pre-coat the plates, diluted antibody (1:40,000 in TBS) was incubated in the wells overnight at 4 °C (50 µL/well) before blocking the remainder of each well with 1% BSA in TBS (200 µL per well) at 37 °C for 1 hour. The reactions added to these pre-prepared plates were incubated at 37°C for 1 hour, before washing 5 times with TBS + 0.1% Tween-20 (200 µL per well). The europium labelled streptavidin (Perkin Elmer) was diluted to 0.1 µg/mL in TBS + 0.1% Tween-20 and added to the wells (50 µL per well) for incubation at 37°C for 30 mins. The plate was washed 5 times with TBS + 0.1% Tween-20 (200 µL per well) and 5 times with

sterile water (200 µL per well). Enhancement solution (Perkin Elmer, 50 µL per well) was added to the wells to incubate for 10 minutes at room temperature before reading the plate using a PerkinElmer Victor X5 multi-label plate reader (time-resolved fluorescence settings, 340 nm excitation and 612 nm emission).

The Michaelis constants for substrates were determined by varying one substrate concentration whilst maintaining constant and saturating concentrations of the remaining substrates. Enzyme activity, measured as a percentage of the maximum activity that was determined when all substrates were present at saturating concentrations, was graphed against the varying substrate concentration using GraphPad Prism. The Michaelis-Menton equation was utilised to apply a non-linear fit to the data to determine the K_M.

Biotinylation assays to determine IC_{50} and K_i of inhibitors

Activity assays were performed as above and according to (30), however, some of the biotin concentrations for the different BPLs were altered to provide saturating conditions (ie 10x K_M). *Sc*BPL and *Zt*BPL both had 15 µM of biotin present, whilst *Bc*BPL had 30 µM. All the BPLs were assayed with *Sc*PC104-GST biotin domain at the saturating concentrations of 15 µM for *Sc*BPL or 20 µM for both *Bc*BPL and *Zt*BPL. The inhibitors were serially diluted to form an 8-point concentration series in 20% DMSO. This series was diluted 1 in 5 into the assay medium, reducing the DMSO concentration to 4%.

The assay data were graphed with the log concentration of inhibitor plotted against the percentage of enzyme activity relative to the no inhibitor control. A sigmoidal curve was fitted using the 'log(inhibitor) vs. response -- Variable slope (four parameters)' function in GraphPad Prism, with the concentration at the point of inflection giving the IC_{50}. This was utilised to calculate the K_i by the following equation:

$$K_i = IC_{50}/((1+[biotin]/K_M(biotin)))$$

Where [biotin] is the concentration of biotin used in the assay, and K_M (biotin) is the Michaelis constant for biotin.

W303 *Saccharomyces cerevisiae* Antimicrobial Assays

Antimicrobial assays to test the susceptibility of *S. cerevisiae* w303 to inhibitors were completed. Inhibitors (at stock concentrations of 2 mg/ml in DMSO) were plated in a 96-well plate to give a 2-fold dilution series in YP media (1% yeast extract, 2% peptone, 2% glucose) with a final DMSO concentration of 3.2% and a starting inhibitor concentration of 64 µg/mL.

Amphotericin B, a licenced antifungal, was included as a positive control (71). An overnight culture of *S. cerevisiae* w303 grown at 30 °C in YP media was diluted to an OD_{600} of 0.01, and 50 µL was used to inoculate the wells of the 96 well plate to give a final starting OD_{600} of 0.005. Plates were grown with agitation at 30 °C for 24 hrs before shaking the plates at 960 rpm on a Multiskan Ascent microtitre plate reader (ThermoFisher Scientific). The OD_{600} was measured using a PerkinElmer Victor X5 multilabel plate reader.

Disc diffusion assays were also completed by growing an overnight culture of *S. cerevisiae* w303 in YP media at 30 °C with agitation at 175 rpm. This was subcultured in fresh YP media to give a starting OD_{600} of 0.1, which was then grown at 30 °C for 3-4 hours until cells were in the mid log phase with an OD_{600} of approximately 0.8. *S. cerevisiae* w303 culture was combined with YP media containing 1.5% agar (2 mL each), and poured onto a YP media with 1.5% agar plate. The overlay was allowed to cool and set before sterile Whatman paper discs (<10 mm in diameter), pre-soaked with 10 µL of 2 mg/mL compound, were placed on the overlay. Plates were left to grow for 4 days at 30 °C then observed for the presence of zones of inhibition around the discs.

S. cerevisiae Viable Cell Measurements

An overnight culture of *S. cerevisiae* w303 in YP media was grown at 30 °C with agitation at 175 rpm. The culture was diluted to an OD_{600} of 0.01, which was serially diluted to a final concentration of 10^{-6}. An equal amount of YP media with 1.5% agar was added to the serially diluted cultures to give a final agar concentration of 0.75% (in 4 mL). This was poured onto YP media with 1.5% agar plates and the overlay allowed to cool and set. Plates were left at 30 °C for 2 days to allow *S. cerevisiae* colonies to form before counting. CFU was determined by the following equation:

$$\text{CFU} = ((\text{number of colonies} \times \text{dilution factor}))/\text{volume plated}$$

Results

Identifying the *bpl* coding regions from *S. cerevisiae, B. cinerea* and *Z. tritici*

The sequence of *Sc*BPL (37,72) has previously been experimentally verified (Supplementary Table 1). However, the *bpl* coding regions in the genomic sequences of *B. cinerea* and *Z. triciti* have not been annotated. BLAST (basic local alignment search tool) was used to predict these coding regions by identifying regions of similarity between these genomes and the *S. cerevisiae bpl1* gene. Sequences were probed for the presence of signature motifs GRGRXG and KWPND that have been previously reported in all BPLs. To confirm the proposed *bpl* genes encoded functional BPLs, the sequences were utilised in a complementation assay in a BPL-temperature sensitive *E. coli* strain (69). Here, the commercially synthesised *bpl* sequences (Geneworks) were sub-cloned into the pAra13 expression vector and transformed into *E. coli* BM4062. This strain possesses a temperature-sensitive *bpl* mutation such that the BPL is active at 30 °C but becomes inactive above 42 °C. At this restrictive temperature, only cells complemented with a recombinant, functional BPL will survive. The class III BPL from the fungi *C. albicans* was included as a positive control, as its sequence has also previously been identified and confirmed for activity using this complementation assay (58). All *E. coli* BM4062 strains tested grew at the permissive temperature of 30 °C (Figure 1A), showing all strains are viable. For the strain harbouring pAra13 no growth was observed at 42 °C (Figure 1B). As expected, the previously verified class III BPLs from *S. cerevisiae* and *C. albicans* complemented growth at 42 °C (Figure 1B). Importantly, the predicted *Bc*BPL and *Zt*BPL sequences also allowed growth at the restrictive temperature (Figure 1B). Therefore, the identified *bpl* sequences within the *B. cinerea* and *Z. tritici* genomes do indeed encode functional BPLs (Supplementary Table 1).

Sequence alignment of class III BPLs

The sequences of the three BPLs studied here were aligned, together with the human BPL sequence previously reported in the literature (Supplementary Figure 1). All sequences possessed the large N-terminal extension characteristic of class III BPLs in addition to the conserved catalytic domain and C-terminal cap. The class III fungal enzymes were moderately similar to the human isoform with approximately 16% to 18% identity and 28% to 30% similarity over the full sequence (Table 1A). However, the fungal enzymes were more similar to each other with between 32% and 52% identity (46% to 67% similarity) (Table 1A). *Bc*BPL and *Zt*BPL were the most homologous enzymes with 51.7% identity.

The BPL sequences were separated into their two domains by first determining the beginning of the catalytic domain, by alignment with the class I BPL from *M. tuberculosis* and the class II BPLs from *E. coli* and *S. aureus*, which each have known structures (Supplementary Figure 1). For example, the boundary in *Sc*BPL was identified as residue P363 – a conserved proline involved in the termination of a secondary structural element in the class I/II enzymes (an α-helix in the class I *Mt*BPL and a β-sheet from the DNA binding domain in the class II *Ec*BPL and *Sa*BPL). The residues prior to P363 form the *Sc*BPL N-terminal extension, and those following make up the catalytic domain. The N-terminal extension was poorly conserved between the fungal enzymes and the human BPL with only 11% to 13% identity (19% to 21% similarity) (Table 1B). Interestingly, the human BPL contains 160 residues on the N-terminus that do not align to the sequences of any other class III BPL. The human enzyme is only approximately 30 residues larger than *Sc*BPL and of equivalent length to *Bc*BPL and *Zt*BPL. Therefore, it was hypothesised that these residues are not extra sequence but rather a rearrangement of the N-terminal extension such that the region of homology, the predicted structured domain, that is located most N-terminally for the fungal enzymes has been moved C-terminally 160 residues in the human enzyme. The linker between this N-terminal domain and the catalytic domain is reduced in the human enzyme compared to the fungal enzymes as a result of this rearrangement. The purpose of this suggested rearrangement is not known, but it does introduce greater sequence divergence to other class III BPLs in this part of the enzyme. Furthermore, the role of these N-terminal 160 residues is not known, but one study has predicted they form an independent domain that interacts with the remainder of the BPL and the biotin domain substrate (73). The N-terminal extension was more highly conserved between different fungal enzymes with 33% to 52% identity (47% to 68% similarity), with the largest level of identity between *Bc*BPL and *Zt*BPL.

The C-terminal catalytic domain responsible for the binding of ligands and the biotinylation reaction is conserved amongst the three fungal BPLs and the human isoenzyme. The catalytic regions of the fungal BPLs (classified as the remainder of the BPL from the N-terminal extension boundary) share 26% to 28% identity (or 43% to 47% similarity) with the catalytic region of the human enzyme (Table 1C). Hence, this region is more conserved amongst the fungal and human enzymes compared to the N-terminal extension. The fungal BPLs demonstrate 33% to 51% identity (46% to 66% similarity) between their catalytic domains, which is similar to the levels of identity observed between their N-terminal extensions. Conservation of the active site residues involved in ligand binding and catalysis were also analysed. The residues from *S. aureus* BPL and *M. tuberculosis* BPL that are

involved in binding the reaction intermediate were identified from their crystal structures (PDB: 3v8l (*Sa*BPL) and 4op0 (*Mt*BPL)) by LigPlot (Supplementary Figure 2, Supplementary Tables 2 and 3), and compared to the corresponding residues of the class III BPLs (Figure 2). Residues identified to interact with the reaction intermediate through hydrogen bonds with their side chains were mostly conserved, including the lysine responsible for aiding catalysis (K138 in *Mt*BPL) and the tryptophan (W74 in *Mt*BPL) involved in key hydrophobic π-π stacking interactions with the purine rings of ATP. However, an arginine residue that contacts the phosphodiester linker of the reaction intermediate in bacterial BPLs is replaced with a conserved asparagine residue in all class III enzymes, resulting in the loss of a positive charge. This is a key difference in the active site between the class I/II and class III enzymes studied here. Residues that form hydrogen bonds to the reaction intermediate via their backbone amide were also well conserved. Even where divergence was evident (residue 38 and 75 in *Mt*BPL), replacement of these residues was not as crucial since the backbone can still contribute to the binding as long as the structure is not perturbed. The residues involved in hydrophobic interactions with the reaction intermediate were more varied, but typically always replaced with similar hydrophobic residues. Overall, the primary sequence analysis suggests the active site is highly conserved amongst the different class III enzymes with some subtle residue differences.

Expression of full length, catalytically active *Sc*BPL

An expression and purification strategy was developed in *S. cerevisiae* to produce sufficient yields of active *Sc*BPL for further studies. Previous recombinant production strategies have involved expression in *E. coli*, however, endogenous proteases cleave the protease susceptible region located between residues 240 and 260, thereby reducing the yield of full length *Sc*BPL (37). Several methods were attempted to produce full length, soluble, active enzyme, including mutating the cleavage site at residue 248 from a lysine to a threonine, expressing an N-terminally c-myc-tagged and C-terminally hexa-hisitidine-tagged *Sc*BPL to selectively purify full length *Sc*BPL, and utilising denaturing lysis conditions before refolding the enzyme. We subsequently identified a yeast expression system was the optimal approach for recombinant protein production. The *S. cerevisiae bpl1* gene was sub-cloned into the pVT100u vector, under control of the constitutive alcohol dehydrogenase promoter, for expression in *S. cerevisiae* w303. This circumvented the endogenous *E. coli* proteases to allow purification of full length *Sc*BPL via immobilised metal ion affinity chromatography and ion exchange chromatography (Figure 3A). A yield of 9 mg *Sc*BPL per litre was achieved at greater than 95% purity after two chromatography steps. The production of full-length, folded

and active protein was verified using mass spectrometry (MS) and enzyme activity assays. Firstly, the molecular mass measured via denatured MS corresponded to the expected molecular mass calculated from the protein sequence plus the addition of 57 Da (*Sc*BPL: calculated 77186 Da v measured 77244 Da) (Table 2, Supplementary Figure 3A), revealing full-length protein was produced. Sequencing of the expression construct confirmed the *Sc*BPL sequence should yield a protein of 77186 Da, therefore, this additional mass may result from post-translational modifications, which are capable of occurring in yeast expression systems. Denatured MS revealed multiple other species with molecular weights greater than 77244 Da, suggesting further post-translational modifications. The identities of these modifications could not be determined from the calculated mass additions. Further LC-MS/MS could be employed to identify and locate these and the 57 Da post-translational modifications. Native nano-electrospray ionisation-MS (nESI-MS) was next employed and revealed the presence of folded protein (Figure 3B), evident by the low charge state distribution of *Sc*BPL present at high m/z values. *Sc*BPL was incubated with biotin and MgATP for these nESI-MS experiments. A native mass measurement of *Sc*BPL revealed a mass addition of approximately 572 Da indicating *Sc*BPL had synthesized and retained the reaction intermediate – demonstrating the protein was active (Table 2). Furthermore, activity assays measured over a time course of 3 hours revealed that *Sc*BPL was active to biotinylate a biotin domain (from *S. cerevisiae)* when incubated with saturating concentrations of all substrates (biotin, MgATP and biotin domain) (Supplementary Figure 4). Therefore, this new expression method within *S. cerevisiae* has allowed the purification of full-length, folded, active *Sc*BPL to higher yields than previously achieved with *E. coli*-based expression systems.

Recombinant production of *Bc*BPL and *Zt*BPL

Whereas *Sc*BPL was successfully produced in its fungal host organism, *Bc*BPL and *Zt*BPL expression was not detected in *S. cerevisiae* using either the pESC-Ura or pVT100u vectors. However, the proteins were successfully produced in *E. coli* following sub-cloning into the pET16b expression vector and transformation into the *E. coli* BL21 (λDE3)-RIPL strain. The two full length BPLs were purified using immobilised metal ion affinity chromatography (IMAC) followed by size exclusion chromatography to achieve >95% purity (Figure 3A). Approximately 4.5 mg of each *Bc*BPL and *Zt*BPL were produced per litre of culture, with both of these enzymes capable of biotinylating a biotin domain from *S. cerevisiae* measured via an *in vitro* biotinylation activity assay (Supplementary Figure 4). Native nESI-MS revealed these proteins were mostly folded with some minor unfolded species present, which were evident

by their high charge states present at *m/z* values below 3000 (Figure 3B). Native MS also confirmed that both enzymes were active, as mass additions that corresponded with the binding of the reaction intermediate were observed. The masses for both proteins measured via denatured MS (Table 2, Supplementary Figure 3B and C) matched those expected from the protein sequences for *Bc*BPL and *Zt*BPL (*Bc*BPL: calculated 79784 Da v measured 79792 Da; *Zt*BPL: calculated 79042 Da v measured 79047 Da). Denatured MS of both enzymes also revealed multiple species with higher molecular weights than expected, suggesting the presence of post-translational modifications for both *Bc*BPL and *Zt*BPL, as for *Sc*BPL. The identities of these post-translational modifications could not be determined via the calculated mass additions and further LC-MS/MS could be employed to investigate the nature and location of these post-translational modifications. Despite this, two novel BPLs were successfully expressed and purified to produce full-length, active enzymes.

All class III BPLs have similar structure and size

The low-resolution structural techniques of circular dichroism (CD) and native ion mobility-MS (IM-MS) were employed to compare the structures of the three fungal enzymes. Currently, there is no atomic resolution structure of a class III BPL. Whilst crystallography attempts are ongoing, CD and IM-MS can provide information about the overall secondary and tertiary/quaternary structure respectively. For both techniques, the three class III BPLs were incubated with saturating concentrations of biotin and MgATP to allow the formation of the reaction intermediate-bound BPLs and, therefore, homogenous solutions of holo-BPL. CD revealed the three fungal BPLs shared similar secondary structures (Figure 4) that contain a large proportion of α-helices, evident by the maxima near 193 nm and two minima at 208 nm and 222 nm that are characteristic of α-helical proteins (74). However, there was the contribution of some β-sheets as the two minima that indicate α-helical structure were not well defined due to the presence of a β-sheet signal at 218 nm. Deconvolution of the spectra to estimate the proportion of helices and β-sheets within each BPL was attempted, however, the high degree of fitting error resulted in non-meaningful secondary structure predictions. Despite this, the spectral shape was ostensibly the same for the three fungal enzymes suggesting they contain conserved secondary structures.

IM-MS was also utilised to assess structural similarities between these three class III BPLs, particularly with respect to their overall globular fold. IM-MS separates molecules based on their mobility through buffer gas (termed drift time) and can inform on the shape and structure of a molecule (75). The charge state distributions for all the folded enzymes were similar, with

ions observed between +15 to +18 for *Sc*BPL, +16 to +20 for *Bc*BPL and +16 to +19 for *Zt*BPL (Figure 3B). The narrow and low charge state distributions indicate the presence of folded protein. The most intense charge state did vary between the different species, with +18 being the most intense for *Bc*BPL, +18 and +17 equally the most intense for *Zt*BPL, and lower charge states of +17 and +16 the most intense for *Sc*BPL. The observation of similar charge states implies that the BPLs are generally structurally similar, however, the differences in the range of and most intense charge states are most likely a reflection of the differences in molecular size of the enzymes. *Bc*BPL and *Zt*BPL are roughly 2 kDa bigger than *Sc*BPL, therefore, assuming they are totally globular, they will have a greater surface area over which they can accumulate more charges. The calculated masses all corresponded to ligand-bound monomeric species of BPL demonstrating all of the BPLs are functional as monomers, able to bind ligands and form the reaction intermediate. In an effort to improve mass resolution in these spectra, higher cone voltages were applied to improve non-specific adduct formation. Whilst low cone voltages were utilised for ionisation to observe the enzymes in their native, ligand-bound states, higher cone voltages of 200 V revealed only partial dissociation of ligand from the enzymes (Figure 3B). These data suggested strong interactions between the BPLs and the reaction intermediate to allow the complex to endure the strong ionisation conditions.

IM data provide information about the shape and size of a molecule. The collision cross section (CCS) can be measured from these data, which reveals the rotationally averaged size of the molecule (75). The average CCS of the three enzymes were similar (*Sc*BPL: 49.8 nm^2; *Bc*BPL: 52.7 nm^2; *Zt*BPL: 50.9 nm^2) (Table 4) with differences within the recognised error limit of the technique (5%) (63). Furthermore, the CCS values for the individual charge states were similar between the three BPLs. This revealed all enzymes have the same rotationally averaged size, and, therefore are likely to have the same overall fold. The CCS distributions of all enzymes for the common charge states of +17 and +18 overlaid, providing further evidence of structural conservation between the enzymes (Figure 5). *Sc*BPL had a slightly narrower CCS distribution compared to *Bc*BPL and *Zt*BPL, as measured by the width at half maximal height and standard deviation (Supplementary Table 4, Figure 5). Therefore, *Sc*BPL was less dynamic and sampled fewer conformations, giving less variation in CCS. The CCS distribution of *Zt*BPL was broader than *Sc*BPL but narrower than *Bc*BPL, suggesting it is more dynamic than *Sc*BPL but that *Bc*BPL was the most mobile of the three BPLs. These IM-MS data suggested there are subtle structural differences between these three enzymes, but no large, overall conformational variability.

To further assess the stability of the BPLs, collision-induced unfolding MS (CIU-MS) experiments (also known as activated IM-MS (aIM-MS)) were completed. This involved collecting ion mobility data across increasing trap collision energies, which promoted the unfolding of proteins prior to entering the drift tube. As the proteins unfold the drift time will increase due to the collision cross section of the ion increasing. Therefore, the different conformations a protein transitions through as it unfolds can be observed, as well as measuring the overall stability of a protein in the gas phase, both of which can indicate structural differences between proteins (76). Unfolding curves were collected up to 200 V, however, the represented data were truncated above 125 V as there was no change (i.e. increase) to the drift time above this voltage (with this maximum drift time assumed to represent completely unfolded protein) and the signal intensity decreased, decreasing the signal-to-noise. All BPLs were revealed to adopt partially folded transition states during their unfolding process, however, the longevity and number of transitions states varied depending on the enzyme (Figure 6). *Sc*BPL passed through two short-lived transition states before completely unfolding (Figure 6). In contrast, *Bc*BPL began unfolding similarly to *Sc*BPL, with one short-lived transition state. However, it then passed through two longer-lived transition states before becoming completely unfolded (Figure 6). *Zt*BPL had the most striking unfolding pattern, passing through 3 to 4 well-defined transitions (depending on charge state). Some of these transitions were not necessarily longer lived than those of the other BPLs but appeared better defined indicating definitive transition state structures due to the presence of the obvious changes in drift time in the unfolding curve (Figure 6). The differences in these unfolding curves reveals there are structural differences between these fungal BPLs contributing to their varied stabilities and unfolding pathways.

The stability of the fungal enzymes was further assessed by comparing the voltages required to begin denaturing and completely unfold each enzyme in the CIU-MS experiments. *Sc*BPL appeared less stable, both beginning to unfold and being completely unfolded at lower voltages than *Bc*BPL or *Zt*BPL (comparing across the 17+ charge state that is common to all three BPLs) (Figure 6 and Supplementary Figure 5a,b). The later two enzymes were more similar beginning to unfold at similar voltages (Figure 6 and Supplementary Figure 5c). *Zt*BPL had a delay in its initial unfolding compared to *Bc*BPL, but *Bc*BPL required higher voltages to reach its most extended conformation compared to *Zt*BPL (Figure 6 and Supplementary Figure 5c). Therefore, these fungal BPLs have different stabilities and unfolding pathways that suggest structural variance, with *Bc*BPL and *Zt*BPL being more stable to gas phase denaturation.

The stability of the three fungal enzymes was corroborated by thermal denaturation assays that measured the melting temperatures of each protein. This measure of solution thermal stability complements the measurements in the gas phase. Melting temperatures were measured in 100 mM ammonium acetate in the presence of ligands biotin and MgATP, reproducing the conditions utilised for the gas-phase stability measurements. Here, divergence was observed in the stability of the three enzymes in solution. *Sc*BPL had a melting temperature of 51.9 °C ± 0.4 °C (Supplementary Figure 6 and Supplementary Table 5), significantly higher than the melting temperatures of *Bc*BPL and *Zt*BPL, which were similar with values of 48.3 °C ± 0.1 °C and 48.9 °C ± 0.3 °C respectively (Supplementary Figure 6 and Supplementary Table 5). These data demonstrate different stabilities for the fungal class III BPLs, with *Sc*BPL being more thermally stable in solution compared to *Bc*BPL and *Zt*BPL. The varied stabilities indicate subtle structural differences between *Sc*BPL compared to *Bc*BPL and *Zt*BPL. It is noteworthy that these solution stability measurements are contrary to the gas phase stability measurements where *Bc*BPL and *Zt*BPL appear more stable to endure higher collision voltages.

Measuring the Michaelis constants for substrates

All of the BPLs produced in this study were confirmed to biotinylate the BCCD from *S. cerevisiae* pyruvate carboxylase at the time of purification. Time course assays were utilised to identify the enzyme concentration and length of time that allowed enzyme activity to be measured in the linear range of the reaction. The conditions optimal for *Sc*BPL were 4.5 nM for 15 minutes, whilst *Bc*BPL and *Zt*BPL required a concentration of 35 nM for 15 minutes and 20 minutes respectively (Supplementary Figure 4).

Quantitative enzyme assays were carried out to determine the Michaelis constants of the three BPLs for the ligands biotin and MgATP. *Sc*BPL had the highest affinity for biotin (K_M 1.4 ± 0.1 µM), with the two other homologs between two-fold (*Zt*BPL K_M 2.7 ± 0.1 µM) and 4-fold (*Bc*BPL K_M 6.2 ± 0.2 µM) less potent (Table 4). In contrast, *Sc*BPL had 2-fold weaker affinity for MgATP (K_M MgATP 29.8 ± 2.5 µM) than either *Bc*BPL (K_M 11.6 ± 1.6 µM) or *Zt*BPL (K_M 11.2 ± 0.9 µM). Hence, the three enzymes demonstrated subtle variances in the K_M values for the substrates biotin and MgATP.

The Michaelis constants for biotin-accepting protein substrates from different fungal species were also determined and compared to a bacterial substrate. All three BPLs had similar K_M

measurements for the BCCD proteins from fungal pyruvate carboxylases of between 1 and 5 µM (Table 4). *Sc*BPL demonstrated a K_M of 1.2 ± 0.1 µM for its native substrate (the C-terminal 104 residues of the pyruvate carboxylase isoform 1). *Sc*BPL also demonstrated a similarly low K_M of 1.2 ± 0.2 µM for the substrate from *C. albicans* (the C-terminal 115 residues of pyruvate carboxylase isoform 1). However, when *Sc*BPL activity was assayed with a substrate from the bacteria *S. aureus* (the C-terminal 90 residues of *S. aureus* pyruvate carboxylase), the K_M value for this biotin domain was approximately 13-fold weaker (K_M (GST-*Sa*PC90): 13 ± 0.5 µM). This demonstrates that *Sc*BPL prefers the fungal substrates.

The Michaelis constants of *Bc*BPL and *Zt*BPL for the fungal biotin domains from *S. cerevisiae* (GST-*Sc*PC104), *C. albicans* (GST-*Ca*PC115) and *B. cinerea* (GST-*Bc*PC103), and the bacterial biotin domain from *S. aureus* (GST-*Sa*PC90) were measured (Table 4). *Bc*BPL had a MIchaelis constant of 5 µM for GST-*Bc*PC103. This was higher than the Michealis constant of *Sc*BPL for its native substrate (K_M (GST-*Sc*PC104): 1.2 ± 0.1 µM). *Bc*BPL had a similar Michaelis constant for the fungal biotin domain from *S. cerevisiae* (K_M (GST-*Sc*PC104): 4.3 ± 0.4 µM) as for its native substrate, and an approximately 2-fold lower K_M for the *C. albicans* biotin domain (K_M (GST-*Ca*PC115): 2.8 ± 0.2 µM). *Zt*BPL had similar K_M values as *Bc*BPL for the *S. cerevisiae* (K_M (GST-*Sc*PC104): 3.9 ± 0.3 µM) and *C. albicans* biotin domains (K_M (GST-*Ca*PC115): 4.4 ± 0.5 µM). The Michaelis constants of *Bc*BPL and *Zt*BPL for the two fungal biotin domains from *S. cerevisiae* and *C. albicans* were roughly 2- to 3.7-fold higher than the *Sc*BPL K_M values for these domains. However, the K_M of *Zt*BPL for the fungal biotin domain from *B. cinerea* was lower (K_M (GST-*Bc*PC103): 1.8 ± 0.4 µM) and more similar to the K_M values of *Sc*BPL for the fungal biotin domains. Interestingly, *Bc*BPL and *Zt*BPL were able to more efficiently utilise the bacterial *S. aureus* biotin domain than *Sc*BPL. *Bc*BPL displayed a similar K_M for this domain as the fungal domains (*Bc*BPL K_M (GST-*Sa*PC90): 3.4 ± 0.1 µM), whilst the *S. aureus* biotin domain was the preferred substrate for *Zt*BPL with a K_M of 1.1 ± 0.1 µM, which was similar to the K_M for the *B. cinerea* biotin domain and approximately 4-fold lower than its K_M for either the *S. cerevisiae* or *C. albicans* biotin domains (Table 4). Overall, class III BPLs display differences in their preferences for the species of biotin domain they biotinylate.

Inhibition of class III enzymes for the design of antifungals

Biotin analogues and bisubstrate ligands that mimic the reaction intermediate, biotinyl-5′-AMP, have previously been exploited as chemical inhibitors against various bacterial BPLs (class I and II) (21,30,32,33,42-44). A panel of these inhibitors (Figure 7) were tested for

activity against the three class III enzymes characterised here, to investigate whether these BPLs could be inhibited *in vitro* (Table 5). The biotin analogue, biotin acetylene (42), modestly inhibited all three class III BPLs with K_i values of 2.3 ± 0.05 µM, 2.8 ± 0.08 µM, and 3.0 ± 0.07 µM for *Sc*BPL, *Zt*BPL, and *Bc*BPL respectively. The reaction intermediate mimic inhibitors, biotinol-5´-AMP (21,24,43,77-79) and BPL199 (manuscript in preparation & Appendix 2), were generally more potent inhibitors with nano-molar K_i values against the fungal BPLs. Biotinol-5´-AMP was most potent against *Zt*BPL, with a K_i of 63 ± 1 nM, and also had strong inhibitory activity against the other BPLs (*Sc*BPL K_i: 186 ± 8 nM; *Bc*BPL K_i: 659 ± 53 nM). BPL199 was the highest affinity inhibitor for *Sc*BPL with a low nM K_i value of 20 ± 2 nM. This inhibitor also showed potent inhibition of *Zt*BPL (K_i: 178 ± 14 nM) but was approximately 3-fold weaker than biotinol-5´-AMP for this enzyme. Interesting, *Bc*BPL was very poorly inhibited by BPL199, with concentrations above 30 µM required to reduce enzyme activity.

Other inhibitors that mimic the reaction intermediate but contain triazole linkers between the biotinyl and adenyl moieties, rather than the non-hydrolysable phosphodiester in biotinol-5′-AMP, which mimics the natural linker from biotinyl-5′-AMP, and the sulfonyl-based linker of BPL199, were also tested. These inhibitors also had different adenine modifications (BPL68 (21), BPL178 (44) and BPL223 (30)) (Figure 7) (Table 5). Overall, these compounds showed mostly weak to no inhibition (even up to inhibitor concentrations as high as 150 µM to 200 µM) against all three enzymes. However, *Bc*BPL and *Zt*BPL were both inhibited by BPL223 with K_i values of 3.9 ± 0.2 µM and 4 ± 0.4 µM respectively, while *Sc*BPL was unaffected.

All six inhibitors (Figure 7) were tested for whole cell activity against the w303 *S. cerevisiae* strain, both in a broth dilution assay and filter disc assay. None of the inhibitors reduced the growth of *S. cerevisiae* in broth dilution assays at concentrations up to 64 µg/ml when seeded with 1.4×10^4 cells (Supplementary Figure 8A). Due to the lack of whole cell activity, the most potent *in vitro Sc*BPL inhibitor, BPL199, was tested in synergy with a known antifungal to see whether it could improve anti-fungal activity. The licensed antifungal amphotericin B was employed for these synergistic studies as its known mechanism of action involves disrupting cell membranes (80,81). Therefore, it was hypothesised that if fungal cell permeability was limiting the activity of the BPL inhibitors, then using a synergistic antifungal that weakened the cell membrane may allow greater penetrance of BPL199 into the cell. However, no additional anti-fungal effect was evident when BPL199 was used together with amphotericin B (Supplementary Figure 8B). A disc diffusion assay was also utilised to determine if all six BPL inhibitors (Figure 7) demonstrated whole cell activity. This assay allows the application of

inhibitors at higher concentrations directly onto a spread plate to determine whether the inhibitor prevented the growth of the fungi around the inhibitor application site (termed a zone of inhibition). None of the inhibitors produced zones of inhibition when filter discs pre-soaked with the inhibitors at higher concentrations of 2 mg/ml were placed on soft agar overlays of *S. cerevisiae* W303 prior to growth (data not shown). Hence, small molecule inhibitors can inhibit the class III BPLs *in vitro*, however, medicinal chemistry and further optimisation will be required to produce therapeutics with whole cell fungal activity.

Discussion

Eukaryotic class III BPLs still remain largely uncharacterised despite the wealth of knowledge regarding bacterial class I and II BPLs. Here we have recombinantly produced and begun the characterisation of three different class III BPLs, namely those from *S. cerevisiae, B. cinerea* and *Z. triciti*. Expression and purification strategies have been optimised to ensure the production of full length, folded, and active *Sc*BPL, *Bc*BPL and *Zt*BPL to greater than 95% purity. These enzymes were expressed in sufficient yields for structural biology and biophysical characterisation.

The class III BPLs are structurally unique from other characterised BPLs due to the presence of a large N-terminal extension. The absence of an atomic resolution structure hampers our understanding of the structure and function of this extension, as well as that of the enzyme as a whole. Lower-resolution structural techniques have been employed here to analyse the overall fold and secondary structural elements of the three class III BPLs studied. In particular, IM-MS and CD have revealed that both the overall globular folds and the secondary structure of all three enzymes are relatively similar. This conserved fold was revealed by CD to be predominately α-helical with some contributing β-sheets. Small differences in the width of the CCS distributions suggested small structural and dynamic variations between the three enzymes. The stability of the three enzymes was compared to further investigate their structural similarity. This was completed in both the gas phase through CIU-MS, and in-solution by thermal denaturation. These techniques revealed differences in the stability and unfolding transition states of each enzyme, which can be attributed to structural differences between the three BPLs. These differences are predicted to be due to subtle differences in the overall structures. Interestingly, the stability of the enzymes differed when comparing a technique measuring thermal stability in solution, and a gas-phase technique using increasing voltages to denature the protein. Whereas *Sc*BPL was more thermally stable, *Bc*BPL and *Zt*BPL endured higher voltages before beginning to unfold. CIU-MS denatures proteins by increasing the internal energy of the ions through collisions. This is similar to thermal denaturation where heat is used to increase the internal energy of the ions. However, in CIU-MS unfolding is prompted as the protein resolves Coulombic frustration. This mechanism of unfolding is different to that of thermal heating, which is performed in a hydrated environment compared to the gas phase utilised in MS. These differences in the unfolding mechanisms can explain the variability in the stability measures between these two techniques. Despite the modest sequence similarities between the fungal enzymes, it is expected that the overall fold of the enzymes is similar to allow them to carry out their

conserved functions of 1) binding biotin and ATP, 2) synthesizing the reaction intermediate and 3) forming a complex with a suitable protein subtrate and 4) transferring the biotinyl group from the reaction intermediate onto the protein substrate. However, subtle structural differences, whether primary sequence or conformational differences, have been identified between the enzymes and these may be responsible for the variability in stability, and also the differences in substrate and inhibitor affinity that were identified.

The overall structural similarity of these different class III BPLs suggests they should have similar binding affinities for the substrates. Activity assays were employed to determine the Michealis constants of each enzyme for its substrates biotin, MgATP, and the biotin domains from different species. These constants were previously determined for *Sc*BPL, albeit utilising an alternative biochemical assay that measures the incorporation of radiolabelled biotin onto protein (37). Most K_M values measured here correspond well with these published values, however, a difference in the K_M for biotin was observed. Whilst previous studies reported a K_M of 67 nM, here, a K_M of 1.4 µM was determined. This is closer to the K_M values for biotin observed for BPLs from other species (*S. aureus:* 1 µM (70), Human: 0.8 µM (82), *E. coli*: 0.3 µM (83), *M. tuberculosis*: 0.42 µM (23)), and for the other BPLs in this study, which had K_M values of 2.7 µM and 6.2 µM for *Zt*BPL and *Bc*BPL respectively. The Michealis constants for MgATP were similar, with only 2- to 3-fold difference between the three fungal BPLs studied. They were also similar to K_M values reported for other BPLs (Human: 47 µM (82), *M. tuberculosis*: 21 µM (23), *S. aureus:* 180 µM (70), *E. coli*: 200 µM (83)). Therefore, whilst the structures of these three enzymes are similar overall, there are subtle differences within the substrate pockets leading to differences in the affinities of the substrates. The sequence alignment reveals only slight sequence divergence in the active site residues involved in binding the reaction intermediate, which we propose may be responsible for these slight substrate affinity differences. However, these sequence differences may provide the opportunity to selectively target one specific fungal BPL for the design of narrow spectrum antifungal treatments, a possibility confirmed by the results of the inhibitor binding assays (*vide infra*).

BCCD protein substrates from *S. cerevisiae, C. albicans, B. cinerea* and *S. aureus* were investigated with the three fungal BPLs. Differences in the Michaelis constants for these biotin domains were observed between the three BPLs. *Sc*BPL had an equal preference for the *S. cerevisiae* and *C. albicans* biotin domains, whilst its affinity for the bacterial biotin domain from *S. aureus* was ten-fold weaker. This is similar to the finding that the Class III human BPL

has a six-fold increase in K_M for a bacterial biotin domain from *E. coli* compared to a human biotin domain (human BPL K_M for *E. coli* biotin domain: 65 µM, K_M for human biotin domain: 11 µM) (84). Previous studies have also shown that the K_M of *Sc*BPL for a biotin domain from its own species is 1.0 ± 0.2 µM (37), which corresponds to the value reported here. This previous study also determined that *Sc*BPL could biotinylate a biotin domain from *E. coli* with a 10-fold higher K_M of 11.1 ± 1 µM (37), similar to the K_M for the *S. aureus* biotin domain reported here. Despite both the *E. coli* and *S. aureus* biotin domains being bacterial, they are notably different structurally. The *E. coli* domain (from acetyl-coA carboxylase, the only *E. coli* biotin-dependent enzyme) contains a protruding loop known as a 'thumb' structure (85,86), which is absent in all other biotin domains identified to date. The similar affinities for these differently structured domains suggests the thumb is not involved in the interaction with *Sc*BPL, and is not responsible for the reduction in affinity between the fungal and bacterial biotin domains.

*Bc*BPL and *Zt*BPL revealed different patterns of biotin domain preferences compared to *Sc*BPL. *Bc*BPL demonstrated similar affinities for all four biotin domains, with no clear preference for its native substrate, while *Zt*BPL showed a preference for the bacterial *S. aureus* and the fungal *B. cinerea* biotin domains over the other two fungal substrates. Therefore, class III enzymes have preferences for the biotin domains they will biotinylate that may be due to undetermined structural features of the biotin domain or BPL, or the differences at the amino acid level between these two proteins. It has been hypothesised in the literature that class III BPLs are selective towards their targets and contain substrate verification activity, however, the mechanistic basis of this is yet to be determined (25,39,84,87-89). The interaction interface of a class III BPL with a biotin domain has not been empirically determined. Therefore, the residues critical for this interaction to allow biotinylation have not been identified and the conservation of these residues between BPLs and biotin domains cannot be elucidated. Subtle structural differences at this protein-protein interaction interface, which may be responsible for substrate selection preferences, could also facilitate the selective inhibition of a specific fungal BPL by small molecules targeting the interface between the BPL and biotin domain to prevent biotinylation.

Several BPL inhibitors active against the *S. aureus* BPL were tested for activity against these three BPLs. These inhibitors were rationally designed as biotin and reaction-intermediate mimics and, hence, some of them were cross-reactive for the fungal BPLs tested here. Nano-molar inhibition constants were determined for the reaction-intermediate mimics biotinol-5′-

AMP and BPL199 for all BPLs, except *Bc*BPL where inhibition by BPL199 was only evident at micro-molar concentrations. These inhibition constants were only modestly weaker (approximately 10- to 30-fold) compared to the inhibition constants of these inhibitors against their intended target *Sa*BPL (biotinol-5′-AMP: K_i 18 nM (77), BPL199: K_i 2.4 nM (manuscript in preparation & Appendix 2)). Therefore, the class III enzymes can be chemically inhibited *in vitro* by direct occupation of both the biotin and MgATP-binding pockets of the active site. Weak inhibition was possible by solely occupying the biotin-binding pocket, as seen with the biotin acetylene inhibitor. However, linking the biotin to an ATP analogue via a favourable chemical linker dramatically increases inhibition of the BPL. These components must be optimal, however, as inhibitors BPL068, BPL178 and BPL223, which contained a triazole linker and different adenine modifications, have inhibition constants weaker or similar to those of biotin acetylene, or do not inhibit enzyme activity at all. Despite only small divergence in active site residues between the human and different fungal BPLs, selective inhibition was demonstrated. *Sc*BPL and *Zt*BPL were inhibited by BPL199, whilst this inhibitor had no effect on the human enzyme (manuscript in preparation & Appendix 2) and very weak inhibition of *Bc*BPL (with complete inhibition not achieved at a final inhibitor concentration of 150 µM, the maximum concentration possible to test, preventing IC_{50} and K_i determination). Therefore, BPL199 was 56-fold more selective for *Zt*BPL and 500-fold selective for *Sc*BPL over *Bc*BPL or human BPL. Biotinol-5′-AMP also demonstrated 10-fold selectivity between *Bc*BPL and *Zt*BPL, with respective K_i values of 659 nM and 63 nM. These two enzymes demonstrated the greatest sequence identity (51.7%) but despite this, two examples of selective inhibition between these enzymes have been demonstrated here. These data demonstrate that selective inhibition between different fungal BPLs to create a narrow spectrum antifungal, without inhibiting the human BPL, is possible despite the highly conserved sequence similarity.

Despite inhibiting the enzymes *in vitro,* none of these compounds were able to inhibit the growth of, or kill, *S. cerevisiae*. As BPL is a cytosolic protein, inhibitors must traverse the complex, protective fungal cell wall, followed by its cell membrane, before being able to effectively act upon their target. Hence, the lack of whole cell efficacy may result from an inability of the inhibitors to penetrate into the fungal cells. Alternatively, the inhibitors may enter the cell but are efficiently exported out again via efflux transport mechanisms. Despite this, the *in vitro* enzyme inhibition demonstrates these BPLs can be selectively inhibited. Structural studies are required to confirm the inhibitors are binding to the active site substrate pockets, and also provide information for the optimisation of inhibitors via structure-guided

medicinal chemistry. Further studies into the uptake or transport of the inhibitors into fungal cells, and also potential efflux pathways, are required to determine the mechanism resulting in the lack of whole cell efficacy. Inhibitors can then be optimised to improve *in vitro* inhibition in combination with whole cell activity. Genomic studies also support BPL as a viable antifungal target as it is an essential enzyme in *S. cerevisiae*, *C. albicans* and *Schizosaccharomyces pombe* (45-48), therefore, it is presumed to be an essential enzyme in all fungal pathogens. BPL knockout studies in the specific fungal organisms to be targeted by BPL inhibitors should be generated to validate BPL as an appropriate target for these pathogens.

The three fungal enzymes studied here were all determined to be moderately similar to each other based on their sequence homology. Particularly, *Bc*BPL and *Zt*BPL were identified as the most similar. With the data presented here, *Bc*BPL and *Zt*BPL also appear to behave more alike, both structurally and kinetically, compared to *Sc*BPL. *Bc*BPL and *Zt*BPL had similar MS charge state distributions, CCS values, solution and gas phase stabilities, unfolding transitions states, and Michaelis and inhibition constants, with similar K_M values for MgATP and being inhibited by a fluorine tri-substituted triazole-linked reaction intermediate mimic (BPL223). These patterns of structure and activity similarity between the fungal enzymes are not surprising based on the sequence alignment analysis.

The class III BPLs from the fungal species *S. cerevisiae, B. cinerea* and *Z. triciti* were recombinantly produced and characterised to provide biochemical and structural insights into this unique structural class of enzymes as they represent novel antifungal targets. IM-MS and CD revealed similar overall structures of the three enzymes as predicted by the high levels of sequence similarity between these BPLs. However, the variable Michaelis constants for the substrates and the differences in the stabilities of these enzymes imply subtle structural differences. These class III BPLs demonstrated different patterns of preference for fungal and bacterial biotin domains, further validating a substrate selection activity in class III BPLs. Finally, potent, selective inhibition of all three enzymes was demonstrated *in vitro*, with BPL199 showing 56-fold selectivity for *Zt*BPL over *Bc*BPL, the two most sequence identical and similarly behaved enzymes identified in this study. Higher-resolution structural information will be required to aid future drug discovery efforts, which can be investigated using the three different class III enzymes now at our disposal. Structural insights into these enzymes will also help answer fundamental questions about eukaryotic class III BPLs, to answer the many questions surrounding their biology that currently remain unanswered.

Acknowledgements

The authors would like to thank Dr Rietie Venter (University of South Australia) for providing amphotericin B for use in the antifungal cell susceptibility assays and expertise for these assays. Circular dichroism experiments were completed at the Biophysical Characterisation Facility at the University of South Australia Cancer Research Institute.

Funding

This work was supported by Bayer Crop Sciences, Grants for Targets programme awarded to SWP. LMS is the recipient of an Australian Government Research Training Program Stipend Scholarship and the Fred Collins Fellowship from the Australian Society of Biochemistry and Molecular Biology.

Author contributions

LMS, GWB, SWP conceived the project and designed the study; LMS, SN, NRP, KP, TLP collected the data; LMS, SN, NRP, KP, TLP, GWB, KLW, SWP analysed the data; MCW, KP, PB, TLP, GWB, KLW, SWP advised on and/or supervised the project; and LMS, TLP, GWB, KLW, SWP prepared the manuscript.

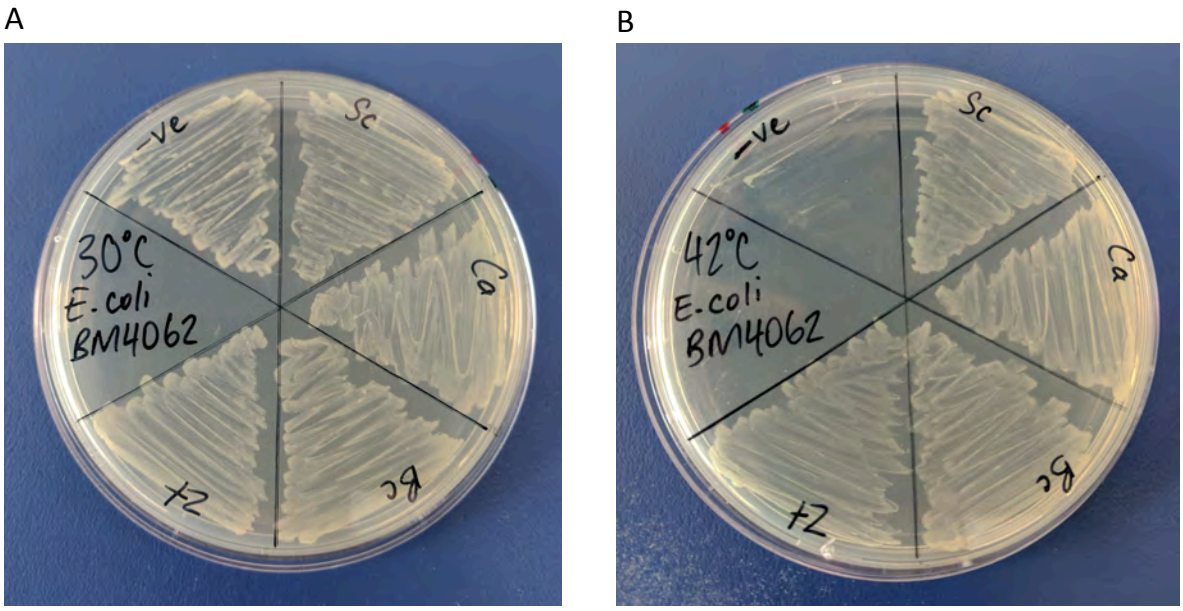


Figure 1. The sequences for the three class III BPLs from *S. cerevisiae, B. cinerea* and *Z. triciti* encode functional BPLs that can rescue mutant *E. coli* strains with a non-functional endogenous BPL when grown at 42 °C. *E. coli* BM4062 were transformed with pAra13 (-ve), pAra13-*Sc*BPL (Sc), pAra13-*Ca*BPL (Ca), pAra13-*Bc*BPL (Bc) and pAra13-*Zt*BPL (Zt), and grown at either A) 30 °C or B) 42 °C. The class III BPL from *C. albicans* was included as a positive control as it has previously been characterised for activity in this complementation assay.

Table 1. Comparison of the percentage sequence identity and similarity (in brackets) of the class III BPLs from humans, *S. cerevisiae*, *B. cinerea* and *Z. tritici* for the A) full length enzymes, B) N-terminal extension, C) catalytic domain and D) active site (classified as residues involved in the binding of the reaction intermediate). The boundary of the N-terminal extension and catalytic domain was determined from alignment with the *M. tuberculosis, S. aureus* and *E. coli* BPLs, and using their crystal structures to determine where their catalytic domains initiate.

A

Full-length Sequence	Human	*Sc*BPL	*Bc*BPL	*Zt*BPL
Human				
*Sc*BPL	16.57 (28.49)			
*Bc*BPL	17.22 (29.85)	34.56 (49.25)		
*Zt*BPL	18.04 (29.45)	32.79 (46.42)	51.65 (67.05)	

B

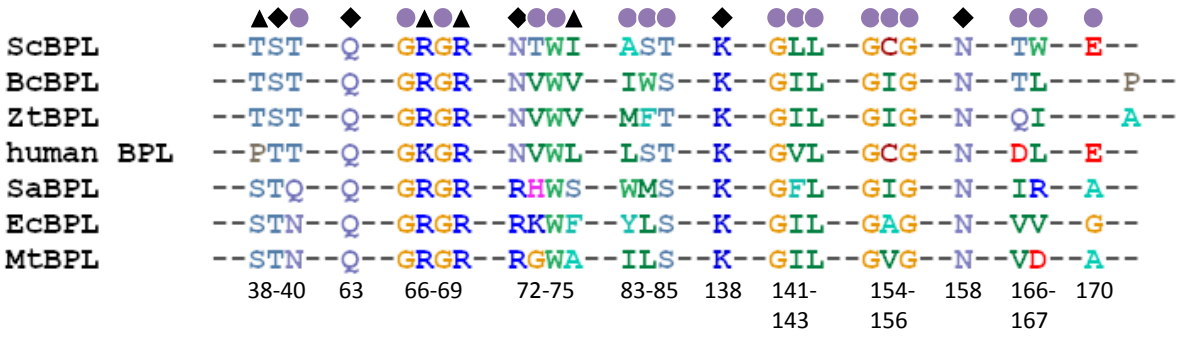
N-terminal Extension	Human	*Sc*BPL	*Bc*BPL	*Zt*BPL
Human				
*Sc*BPL	10.64 (19.82)			
*Bc*BPL	11.43 (20.36)	34.67 (48.99)		
*Zt*BPL	12.59 (20.68)	32.83 (46.46)	51.81 (68.13)	

C

Catalytic Domain	Human	*Sc*BPL	*Bc*BPL	*Zt*BPL
Human				
*Sc*BPL	26.19 (42.56)			
*Bc*BPL	27.65 (46.95)	34.42 (49.55)		
*Zt*BPL	27.50 (44.69)	32.75 (46.38)	51.46 (65.70)	

D

Active Site	Human	*Sc*BPL	*Bc*BPL	*Zt*BPL
Human				
*Sc*BPL	58.62 (86.21)			
*Bc*BPL	50.00 (70.00)	56.67 (70.00)		
*Zt*BPL	53.33 (76.67)	56.67 (70.00)	72.41 (82.76)	



▲ Hydrogen-bond through residue backbone
◆ Hydrogen-bond through residue side chain
● Hydrophobic interaction

Figure 2. Sequence alignment showing the conservation of the residues involved in binding the reaction intermediate (biotinyl-5′-AMP) in the active site of various BPLs. These residues were identified from co-crystal structures of the *M. tuberculosis* (PDB: 4OP0) (22) and *S. aureus* (PDB: 3RIR) (29) BPLs using LigPlot+ (56,57). Residue numbering (shown below the sequence alignment) corresponds to the *M. tuberculosis* BPL sequence.

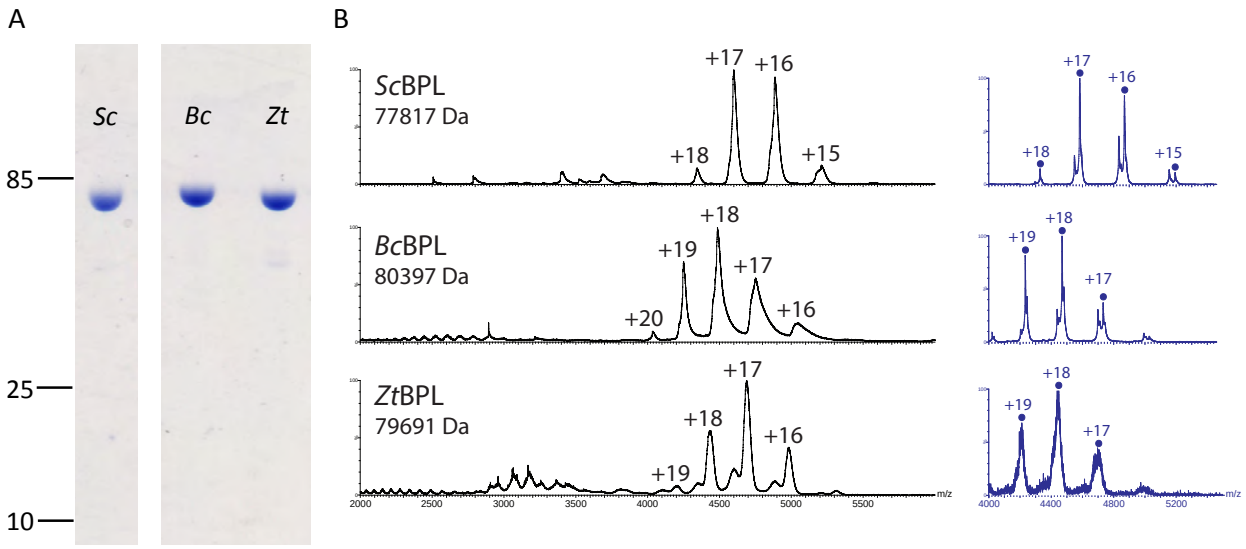


Figure 3. The BPLs from *S. cerevisiae, B. cinerea* and *Z. tritici* were recombinantly expressed and A) purified to near homogeneity (5 µg of each loaded on SDS-PAGE). B) Native nESI-MS was completed to assess the folded state and molecular mass of each *Sc*BPL, *Bc*BPL, and *Zt*BPL in the presence of biotin and MgATP, at cone voltages of 50 V (black) and 200 V (blue). Measured holo-BPL masses are annotated.

Table 2. Summary of the calculated and measured molecular masses of *Sc*BPL, *Bc*BPL and *Zt*BPL. Masses were experimentally determined by both denatured MS and native nESI-MS.

	Apo				**Holo**		
	Calculated Mass from Sequence (Da)	**Measured Denatured Mass (Da)**	**Denatured measured mass - calculated mass (Da)**	**Measured Native Mass (Da)**	**Calculated Mass from Sequence (Da)**	**Calculated Mass from Denatured Measured Mass (Da)**	**Measured Native Mass (Da)**
ScBPL	77186	77244	58	77275	77759	77817	77817
BcBPL	79784	79792	8	79825	80357	80365	80397
ZtBPL	79042	79047	5	79114	79615	79620	79691

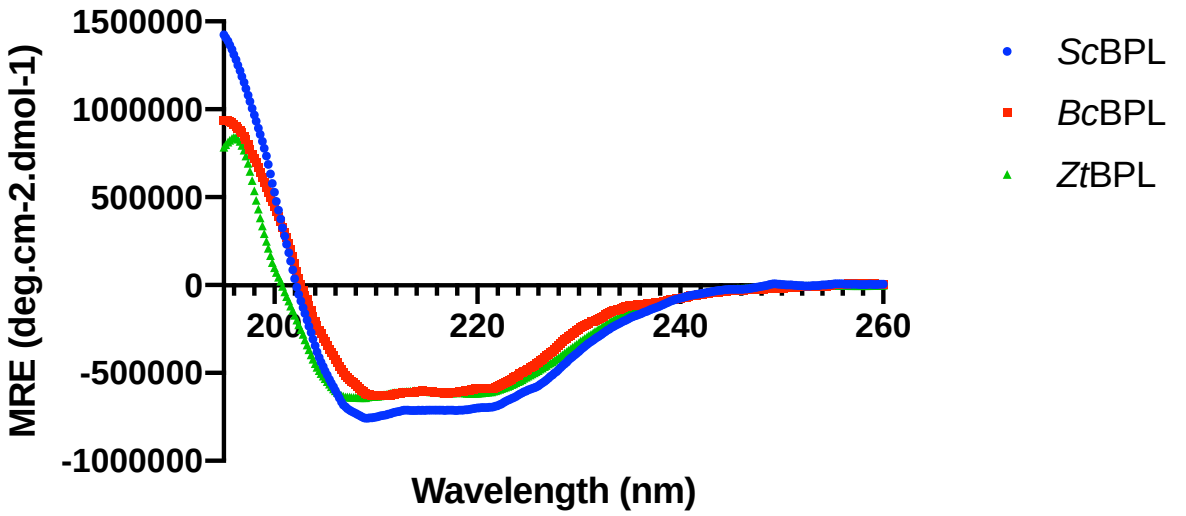


Figure 4. The fungal BPLs from *S. cerevisiae* (blue), *B. cinerea* (red) and *Z. tritici* (green) have similar secondary structure shown by circular dichroism. Spectra are the smoothed average of 5 scans corrected for buffer and protein concentration.

Table 3. Measured collision cross sections (CCS) of the class III fungal BPLs using IM-MS. The associated error for CCS determination is 5% (63).

			CCS (nm^2)						
	Sequence Length	**Molecular Mass**	**15+**	**16+**	**17+**	**18+**	**19+**	**20+**	**Average CCS (nm^2)**
***Sc*BPL**	**696**	**77244**	48.05	48.93	50.11	51.91			49.75
***Bc*BPL**	**723**	**79784**			50.87	51.74	53.27	55.01	52.72
***Zt*BPL**	**731**	**79042**		49.07	49.96	51.39	53.27		50.92

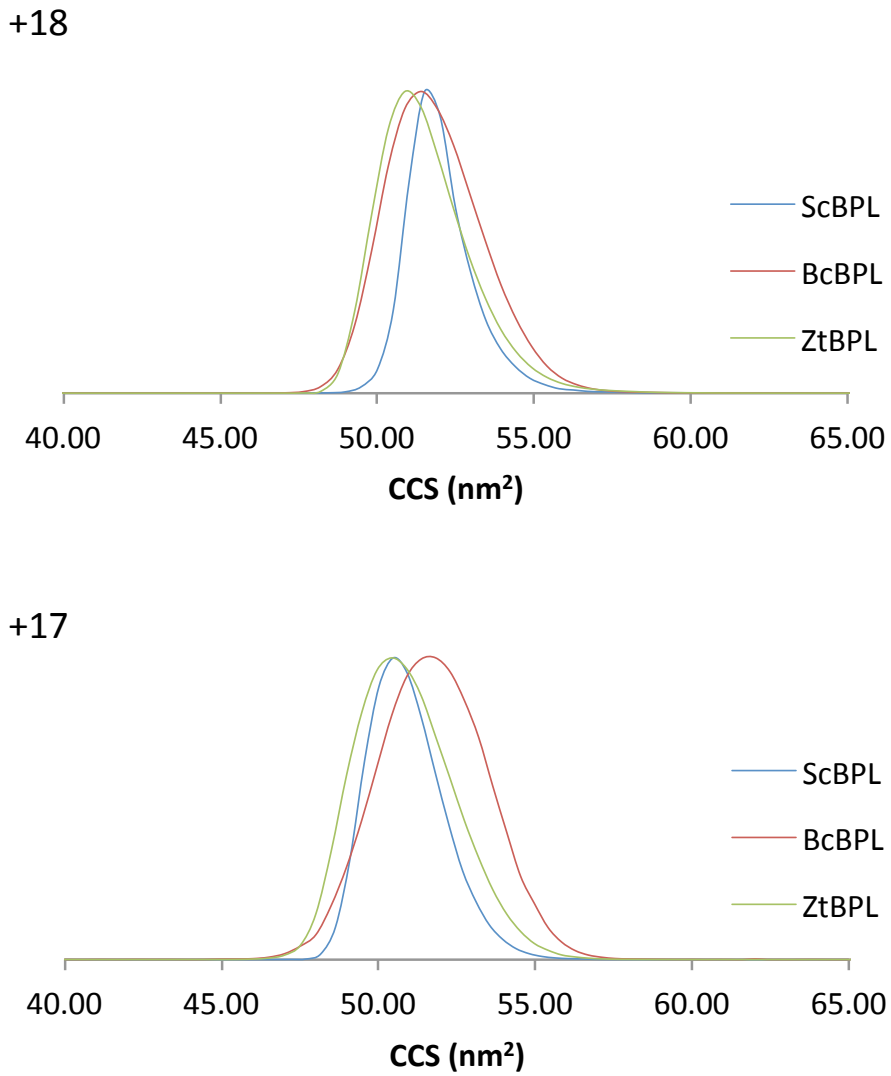


Figure 5. The CCS distributions for the common charge states (+17 and +18) of the class III BPLs *Sc*BPL (blue), *Bc*BPL (red), and *Zt*BPL (green). The associated error for this technique is 5% (63).

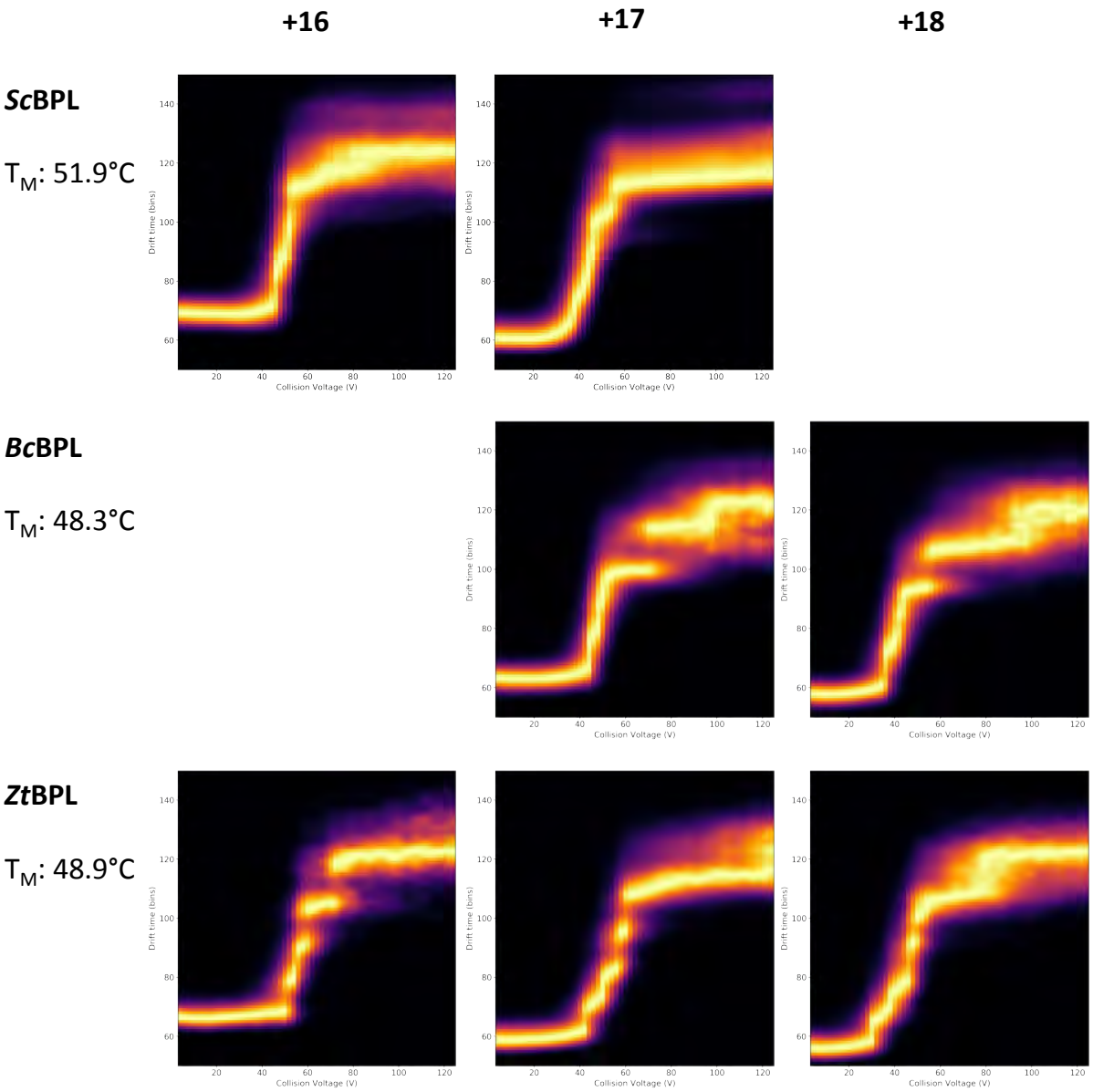


Figure 6. Unfolding curves of the three fungal BPLs measured by CIU-MS, revealing protein stability in the gas phase. Annotated are the measurements of thermal stability of each enzyme, as determined by in-solution thermal denaturation assays.

Table 4. Michaelis constants (K_M) of the three fungal BPLs for biotin, MgATP and the biotin domains from different species. ND = not determined.

	***Sc*BPL**	***Bc*BPL**	***Zt*BPL**
Biotin K_M (μM)	1.4 ± 0.1	6.2 ± 0.2	2.7 ± 0.1
MgATP K_M (μM)	29.8 ± 2.5	11.6 ± 1.6	11.2 ± 0.9
GST-*Sc*PC104 K_M (μM)	1.2 ± 0.1	4.3 ± 0.4	3.9 ± 0.3
GST-*Ca*PC115 K_M (μM)	1.2 ± 0.2	2.8 ± 0.2	4.4 ± 0.5
GST-*Bc*PC103 K_M (μM)	ND	5.0 ± 0.6	1.8 ± 0.4
GST-*Sa*PC90 K_M (μM)	13 ± 0.5	3.7 ± 0.2	1.1 ± 0.1

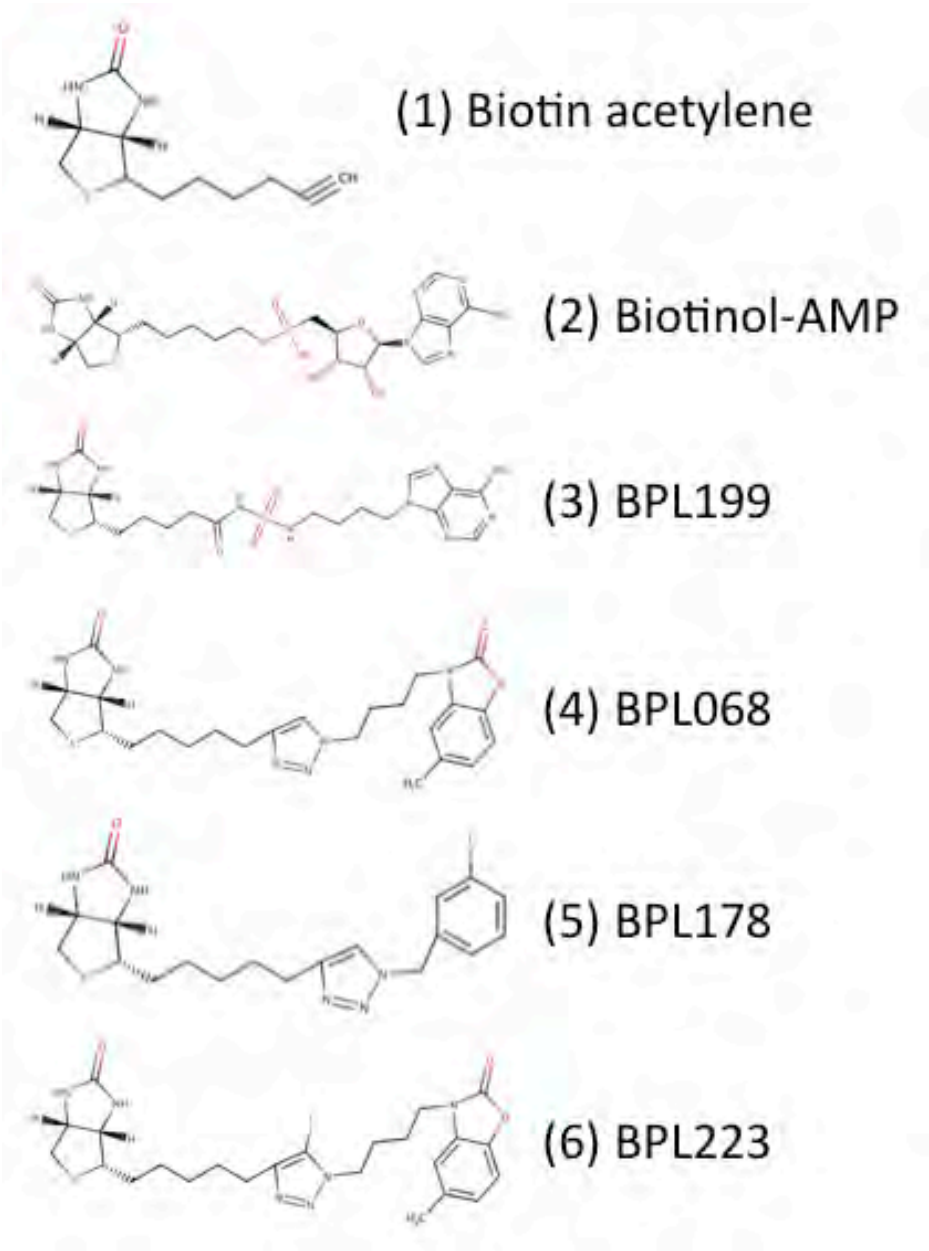


Figure 7. Structures of the BPL inhibitors tested for *in vitro* inhibition of the three fungal enzymes and whole cell antifungal activity against *S. cerevisiae* w303. These were rationally designed inhibitors that mimic the natural substrates biotin and biotinyl-5′-AMP for the inhibition of the *S. aureus* BPL.

Table 5. Inhibition constants of BPL inhibitors against the different fungal BPLs. These inhibitors were rationally designed to mimic the natural substrates biotin and biotinyl-5-AMP, and were used for targeting the BPL from *S. aureus.*

	***Sc*BPL**	***Bc*BPL**	***Zt*BPL**
biotin acetylene K_i (μM)	2.3 ± 0.05	3 ± 0.07	2.8 ± 0.08
biotinol-5-AMP K_i (μM)	0.186 ± 0.008	0.659 ± 0.053	0.063 ± 0.001
BPL199 K_i (μM)	0.020 ± 0.002	> 10	0.178 ± 0.014
BPL068 K_i (μM)	> 10	> 10	> 10
BPL178 K_i (μM)	> 10	> 10	> 10
BPL223 K_i (μM)	> 10	3.9 ± 0.2	4 ± 0.4

References:

1. Bongomin, F., Gago, S., Oladele, R. O., and Denning, D. W. (2017) Global and Multi-National Prevalence of Fungal Diseases - Estimate Precision. *J Fungi* **3**
2. Brown, G. D., Denning, D. W., Gow, N. A. R., Levitz, S. M., Netea, M. G., and White, T. C. (2012) Hidden Killers: Human Fungal Infections. *Sci Transl Med* **4**
3. Veloso, J., and van Kan, J. A. L. (2018) Many Shades of Grey in *Botrytis*-Host Plant Interactions. *Trends Plant Sci* **23**, 613-622
4. Imada, K., Tanaka, S., Ibaraki, Y., Yoshimura, K., and Ito, S. (2014) Antifungal effect of 405-nm light on *Botrytis cinerea*. *Lett Appl Microbiol* **59**, 670-676
5. AbuQamar, S., Moustafa, K., and Tran, L. S. (2017) Mechanisms and strategies of plant defense against *Botrytis cinerea*. *Crit Rev Biotechnol* **37**, 262-274
6. McDonald, M. C., McDonald, B. A., and Solomon, P. S. (2015) Recent advances in the *Zymoseptoria tritici*-wheat interaction: insights from pathogenomics. *Front Plant Sci* **6**
7. McDonald, B. A., and Mundt, C. C. (2016) How Knowledge of Pathogen Population Biology Informs Management of Septoria Tritici Blotch. *Phytopathology* **106**, 948-955
8. Fones, H., and Gurr, S. (2015) The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal Genet Biol* **79**, 3-7
9. O'Driscoll, A., Kildea, S., Doohan, F., Spink, J., and Mullins, E. (2014) The wheat-*Septoria* conflict: a new front opening up? *Trends Plant Sci* **19**, 602-610
10. Chapman, B., Slavin, M., Marriott, D., Halliday, C., Kidd, S., Arthur, I., Bak, N., Heath, C. H., Kennedy, K., Morrissey, C. O., Sorrell, T. C., van Hal, S., Keighley, C., Goeman, E., Underwood, N., Hajkowicz, K., Hofmeyr, A., Leung, M., Macesic, N., Botes, J., Blyth, C., Cooley, L., George, C. R., Kalukottege, P., Kesson, A., McMullan, B., Baird, R., Robson, J., Korman, T. M., Pendle, S., Weeks, K., Liu, E., Cheong, E., and Chen, S. (2017) Changing epidemiology of candidaemia in Australia. *J Antimicrob Chemother* **72**, 1103-1108
11. Leading International Fungal Education. Fungal Infections > Candidaemia and Invasive Candidiasis.
12. Torriani, S. F. F., Melichar, J. P. E., Mills, C., Pain, N., Sierotzki, H., and Courbot, M. (2015) *Zymoseptoria tritici*: A major threat to wheat production, integrated approaches to control. *Fungal Genet Biol* **79**, 8-12
13. Maubon, D., Garnaud, C., Calandra, T., Sanglard, D., and Cornet, M. (2014) Resistance of *Candida* spp. to antifungal drugs in the ICU: where are we now? *Intensive Care Med* **40**, 1241-1255
14. Prasad, R., Banerjee, A., and Shah, A. H. (2017) Resistance to antifungal therapies. *Essays Biochem* **61**, 157-166
15. Kullberg, B. J., and Arendrup, M. C. (2015) Invasive Candidiasis. *N Engl J Med* **373**, 1445-1456
16. Chakrabarti, A. (2011) Drug resistance in fungi - an emerging problem. *Regional Health Forum* **15**, 97-103
17. Tong, L. (2013) Structure and function of biotin-dependent carboxylases. *Cell Mol Life Sci* **70**, 863-891
18. Chapman-Smith, A., and Cronan Jr, J. E. (1999) The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* **24**, 359-363
19. Lane, M. D., Rominger, K. L., Young, D. L., and Lynen, F. (1964) The Enzymatic Synthesis of Holotranscarboxylase from Apotranscarboxylase and (+)-Biotin. *J Biol Chem* **239**, 2865-2871

20. Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2008) Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate. *J Biol Chem* **283**, 14739-14750
21. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Pendini, N. R., Polyak, S. W., Pedersen, D. S., Morona, R., Turnidge, J. D., Wallace, J. C., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2012) Selective inhibition of Biotin Protein Ligase from *Staphylococcus aureus*. *J Biol Chem* **287**, 17823-17832
22. Ma, Q., Akhter, Y., Wilmanns, M., and Ehebauer, M. T. (2014) Active site conformational changes upon reaction intermediate biotinyl-5′ -AMP binding in biotin protein ligase from *Mycobacterium tuberculosis*. *Protein Sci* **23**, 932-939
23. Purushothaman, S., Gupta, G., Srivastava, R., Ramu, V. G., and Surolia, A. (2008) Ligand Specificity of Group 1 Biotin Protein Ligase of *Mycobacterium tuberculosis*. *PLoS ONE* **3**
24. Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D., and Matthews, B. W. (2006) Co-repressor Induced Order and Biotin Repressor Dimerization: A Case for Divergent Followed by Convergent Evolution. *J Mol Biol* **357**, 509-523
25. Mayende, L., Swift, R. D., Bailey, L. M., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., and Polyak, S. W. (2012) A novel molecular mechanism to explain biotin-unresponsive holocarboxylase synthetase deficiency. *J Mol Med* **90**, 81-88
26. Weaver, L. H., Kwon, K., Beckett, D., and Matthews, B. W. (2001) Corepressor-induced organization and assembly of the biotin repressor: A model for allosteric activation of a transcriptional regulator. *PNAS* **98**, 6045-6050
27. Eginton, C., Cressman, W. J., Bachas, S., Wade, H., and Beckett, D. (2015) Allosteric Coupling *via* Distant Disorder-to-Order Transitions. *J Mol Biol* **427**, 1695-1704
28. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., and Matthews, B. W. (1992) *Escherichia coli* biotin holoenzyme synthetase/*bio* repressor crystal structure delineates the biotin- and DNA-binding domains. *PNAS* **89**, 9257-9261
29. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterisation of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci* **22**, 762-773
30. Paparella, A. S., Lee, K. J., Hayes, A. J., Feng, J., Feng, Z., Cini, D., Deshmukh, S., Booker, G. W., Wilce, M. C. J., Polyak, S. W., and Abell, A. D. (2018) Halogenation of Biotin Protein Ligase Inhibitors Improves Whole Cell Activity against *Staphylococcus aureus*. *ACS Infect Dis* **4**, 175-184
31. Gupta, V., Gupta, R. K., Khare, G., Salunke, D. M., Surolia, A., and Tyagi, A. K. (2010) Structural Ordering of Disordered Ligand-Binding Loops of Biotin Protein Ligase into Active Conformations as a Consequence of Dehydration. *PLoS ONE* **5**
32. Duckworth, B. P., Geders, T. W., Tiwari, D., Boshoff, H. I., Sibbald, P. A., Barry III, C. E., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2011) Bisubstrate Adenylation Inhibitors of Biotin Protein Ligase from *Mycobacterium tuberculosis*. *Chem Biol* **18**, 1432-1441
33. Bockman, M. R., Kalinda, A. S., Petrelli, R., De la Mora-Rey, T., Tiwari, D., Liu, F., Dawadi, S., Nandakumar, M., Rhee, K. Y., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2015) Targeting *Mycobacterium tuberculosis* Biotin Protein Ligase (MtBPL) with Nucleoside-Based Bisubstrate Adenylation Inhibitors. *J Med Chem* **58**, 7349-7369
34. Bagautdinov, B., Kuroishi, C., Sugahara, M., and Kunishima, N. (2005) Crystal Structures of Biotin Protein Ligase from *Pyrococcus horikoshii* OT3 and its Complexes: Structural Basis of Biotin Activation. *J Mol Biol* **353**, 322-333
35. Tron, C. M., McNae, I. W., Nutley, M., Clarke, D. J., Cooper, A., Walkinshaw, M. D., Baxter, R. L., and Campopiano, D. J. (2009) Structural and Functional Studies of the

Biotin Protein Ligase from *Aquifex aeolicus* Reveal a Critical Role for a Conserved Residue in Target Specificity *J Mol Biol* **387**, 129-146
36. Satiaputra, J., Shearwin, K. E., Booker, G. W., and Polyak, S. W. (2016) Mechanisms of biotin-regulated gene expression in microbes. *Synth Syst Biotechnol* **1**, 17-24
37. Polyak, S. W., Chapman-Smith, A., Brautigan, P. J., and Wallace, J. C. (1999) Biotin Protein Ligase from *Saccharomyces cerevisiae*. The N-terminal domain is required for complete activity. *J Biol Chem* **274**, 32847-32854
38. Hassan, Y. I., Moriyama, H., Olsen, L. J., Bi, X., and Zempleni, J. (2009) N- and C-terminal domains in human holocarboxylase synthetase participate in substrate recognition. *Mol Genet Metab* **96**, 183-188
39. Sternicki, L. M., Wegener, K. L., Bruning, J. B., Booker, G. W., and Polyak, S. W. (2017) Mechanisms Governing Precise Protein Biotinylation. *Trends Biochem Sci* **42**, 383-394
40. Gerdes, S. Y., Scholle, M. D., Campbell, J. W., Balázsi, g., Ravasz, E., Daugherty, M. D., Somera, A. L., Kyrpides, N. C., Anderson, I., Gelfand, M. S., Bhattacharya, A., Kapatral, V., D'Souza, M., Baev, M. V., Grechkin, Y., Mseeh, F., Fonstein, M. Y., Overbeek, R., Barabási, A.-L., Oltvai, Z. N., and Osterman, A. L. (2003) Experimental Determination and System Level Analysis of Essential Genes in *Escherichia coli* MG1655. *J Bacteriol* **185**, 5673-5684
41. Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40
42. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Zvarec, O., Bell, J. M., Turnidge, J. D., Wallace, J. C., Booker, G. W., Wilce, M. C. J., Abell, A. D., and Polyak, S. W. (2012) Biotin Analogues with Antibacterial Activity Are Potent Inhibitors of Biotin Protein Ligase. *ACS Med Chem Lett* **3**, 509-514
43. Tieu, W., Jarrad, A. M., Paparella, A. S., Keeling, K. A., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., Polyak, S. W., and Abell, A. D. (2014) Heterocyclic acyl-phosphate bioisostere-based inhibitors of *Staphylococcus aureus* biotin protein ligase. *Bioorg Med Chem Lett* **24**, 4689-4693
44. Feng, J., Paparella, A. S., Tieu, W., Heim, D., Clark, S., Hayes, A., Booker, G. W., Polyak, S. W., and Abell, A. D. (2016) New Series of BPL Inhibitors To Probe the Ribose-Binding Pocket of *Staphylococcus aureus* Biotin Proteing Ligase *ACS Med Chem Lett* **7**, 1068-1072
45. Hoja, U., Wellein, C., Greiner, E., and Schweizer, E. (1998) Pleiotropic phenotype of acetyl-CoA-carboxylase-defective yeast cells: Viability of a BPL1-amber mutation depending on its readthrough by normal tRNA (Gln)(CAG). *Eur J Biochem* **254**, 520-526
46. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Rhonda, B., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., Bakkoury, M. E., Foury, F., Friend, S. H., Gentalen, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D. J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J. L., Riles, L., Roberts, C. J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R. K., Véronneau, S., Voet, M., Volckaert, G., Ward, T. R., Wysocki, R., Yen, G. S., Yu, K., Zimmermann, K., and Philippsen, P. (1999) Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science* **285**, 901-906
47. Shtifman Segal, E., Gritsenko, V., Levitan, A., Yadav, B., Dror, N., Steenwyk, J. L., Silberberg, Y., Mielich, K., Rokas, A., Gow, N. A. R., Kunze, R., Sharan, R., and Berman, J. (2018) Gene Essentiality Analyzed by *In Vivo* Transposon Mutagenesis and Machine Learning in a Stable Haploid Isolate of *Candida albicans*. *mBio* **9**

48. Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H.-S., Duhig, T., Nam, M., Palmer, G., Han, S., Jeffery, L., Baek, S.-T., Lee, H., Shim, Y. S., Lee, M., kIM, L., Heo, K.-S., Noh, E. J., Lee, A.-R., Jang, Y.-J., Chung, K.-S., Choi, S.-J., Park, J.-Y., Park, Y., Kim, H. M., Park, S.-K., Park, H.-J., Kang, E.-J., Kim, H. B., Kang, H.-S., Park, H.-M., Kim, K., Song, K., Song, K. B., Nurse, P., and Hoe, K.-L. (2010) Analysis of a genome-wide set of gene deletions in the fissuion yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* **28**, 617-623
49. Haßlacher, M., Ivessa, A. S., Paltauf, F., and Kohlwein, S. D. (1993) Acetyl-CoA Carboxylase from Yeast Is an Essential Enzyme and Is Regulated by Factors That Control Phospholipid Metabolism. *J Biol Chem* **268**, 10946-10952
50. Walker, M. E., Val, D. L., Rohde, M., Devenish, R. J., and Wallace, J. C. (1991) Yeast pyruvate carboxylase: identification of two genes encoding isoenzymes. *Biochem Biophys Res Commun* **176**, 1210-1217
51. Navarathna, D. H. M. L. P., Lionakis, M. S., Lizak, M. J., Munasinghe, J., Nickerson, K. W., and Roberts, D. D. (2012) Urea Amidolyase (*DUR1,2*) Contributes to Virulence and Kidney Pathogenesis of *Candida albicans*. *PLoS One* **7**
52. Sierotzki, H., and Scalliet, G. (2013) A Review of Current Knowledge of Resistance Aspects for the Next-Generation Succinate Dehydrogenase Inhibitor Fungicides. *Phytopathology* **103**, 880-887
53. Stucka, R., Dequin, S., Salmon, J. M., and Gancedo, C. (1991) DNA sequences in chromosomes II and VII code for pyruvate carboxylase isoenzymes in *Saccharomyces cerevisiae*: analysis of pyruvate carboxylase-deficient strains. *Mol Gen Genet* **229**, 307-315
54. Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95-98
55. Stothard, P. (2000) The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *Biotechniques* **28**, 1102-1104
56. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-134
57. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786
58. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C. J., Wallace, J. C., and Polyak, S. W. (2008) Biotin protein ligase from *Candida albicans*: Expression, purification and development of a novel assay. *Arch Biochem Biophys* **479**, 163-169
59. Vernet, T., Dignard, D., and Thomas, D. Y. (1987) A family of yeast expression vectors containing the phage f1 intergenic region. *Gene* **52**, 225-233
60. Rothstein, R. J. (1983) One-step gene disruption in yeast. *Methods Enzymol.* **101**, 202-211
61. Satiaputra, J., Sternicki, L. M., Hayes, A. J., Pukala, T. L., Booker, G. W., Shearwin, K. E., and Polyak, S. W. (2019) Native mass spectrometry identifies an alternate DNA-binding pathway for BirA from *Staphylococcus aureus*. *Scientific Reports* **9**
62. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* (Walker, J. M. ed.), Humana Press. pp 571-607
63. Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V., and Ruotolo, B. T. (2010) Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal Chem* **82**, 9557-9565
64. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc* **3**, 1139-1152

65. Migas, L. G., France, A. P., Bellina, B., and Barran, P. E. (2017) ORIGAMI: A software suite for activated ion mobility mass spectrometry (aIM-MS) applied to multimeric protein assemblies. *Int J Mass Spectrom* **427**, 20-28
66. Boivin, S., Kozak, S., and Meijers, R. (2013) Optimization of protein purification and characterization using Thermofluor screens. *Protein Expr Purif* **91**, 192-206
67. Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J., and Weiss, M. S. (2013) Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**, 209-214
68. Val, D. L., Chapman-Smith, A., Walker, M. E., Cronan Jr, J. E., and Wallace, J. C. (1995) Polymorphism of the yeast pyruvate carboxylase 2 gene and protein: effects on protein biotinylation. *Biochem J* **312**, 817-825
69. Barker, D. F., and Campbell, A. M. (1980) Use of *bio-lac* Fusion Strains to Study Regulation of Biotin Biosynthesis in *Escherichia coli*. *J Bacteriol* **143**, 789-800
70. Soares da Costa, T. P., Yap, M. Y., Perugini, M. A., Wallace, J. C., Abell, A. D., Wilce, M. C. J., Polyak, S. W., and Booker, G. W. (2014) Dual roles of F123 in protein homodimerization and inhibitor binding to biotin protein ligase from *Staphylococcus aureus*. *Mol Microbiol* **91**, 110-120
71. Mesa-Arango, A. C., Scorzoni, L., and Zaragoza, O. (2012) It only takes one to do many jobs: Amphotericin B as antifungal and immunomodulatory drug. *Front Microbiol* **3**
72. Cronan Jr, J. E., and Wallace, J. C. (1995) The gene encoding the biotin-apoprotein ligase of *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* **130**, 221-229
73. Lee, C.-K., Cheong, C., and Jeon, Y. H. (2010) The N-terminal domain of human holocarboxylase synthetase facilitates biotinylation via direct interaction with the substrate protein. *FEBS Lett* **584**, 675-680
74. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876-2890
75. Pukala, T. (2018) Importance of Collision Cross Section Measurements by Ion Mobility-Mass Spectrometry in Structural Biology. *Rapid Commun Mass Spec*
76. Dixit, S. M., Polasky, D. A., and Ruotolo, B. T. (2018) Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Curr Opin Chem Biol* **42**, 93-100
77. Tieu, W., Polyak, S. W., Paparella, A. S., Yap, M. Y., Soares da Costa, T. P., Ng, B., Wang, G., Lumb, R., Bell, J. M., Turnidge, J. D., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2015) Improved Synthesis of Biotinol-5′-AMP: Implications for Antibacterial Discovery. *ACS Med Chem Lett* **6**, 216-220
78. Brown, P. H., Cronan, J. E., Grøtli, M., and Beckett, D. (2004) The Biotin Repressor: Modulation of Allostery by Corepressor Analogs. *J Mol Biol* **337**, 857-869
79. Sittiwong, W., Cordonier, E. L., Zempleni, J., and Dussault, P. H. (2014) β-Keto and β-hydroxyphosphonate analogs of biotin-5-AMP are inhibitors of holocarboxylase synthetase. *Bioorg Med Chem Lett* **24**, 5568-5571
80. Ellis, D. (2002) Amphotericin B: spectrum and resistance. *J Antimicrob Chemother* **49**, 7-10
81. Dutcher, J. D., Gold, W., Pagano, J. F., and Vandepatte, J. (1959) Amphotericin B, its production and its salts. (Office, U. S. P. ed., United States of America
82. Ingaramo, M., and Beckett, D. (2009) Distinct Amino Termini of Two Human HCS Isoforms Influence Biotin Acceptor Substrate Recognition. *J Biol Chem* **284**, 30862-30870
83. Chapman-Smith, A., Mulhern, T. D., Whelan, F., Cronan Jr, J. E., and Wallace, J. C. (2001) The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity. *Protein Sci* **10**, 2608-2617

84. Healy, S., McDonald, M. K., Wu, X., Yue, W. W., Kochan, G., Oppermann, U., and Gravel, R. A. (2010) Structural Impact of Human and *Escherichia coli* Biotin Carboxyl Carrier Proteins on Biotin Attachment. *Biochemistry* **49**, 4687-4694
85. Yao, X., Wei, D., Soden, C., Summers, M. F., and Beckett, D. (1997) Structure of the Carboxy-Terminal Fragment of the Apo-Biotin Carboxyl Carrier Subunit of *Escherichia coli* Acetyl-CoA Carboxylase. *Biochemistry* **36**, 15089-15100
86. Athappilly, F. K., and Hendrickson, W. A. (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**, 1407-1419
87. Ingaramo, M., and Beckett, D. (2011) Biotinylation, a Post-translational Modification Controlled by the Rate of Protein-Protein Association. *J Biol Chem* **286**, 13071-13078
88. Ingaramo, M., and Beckett, D. (2012) Selectivity in Post-translational Biotin Addition to Five Human Carboxylases. *J Biol Chem* **287**, 1813-1822
89. Campeau, E., and Gravel, R. A. (2001) Expression in *Escherichia coli* of N- and C-terminally Deleted Human Holocarboxylase Synthetase: Influence of the N-terminus on biotinylation and identification of a minimum functional protein. *J Biol Chem* **276**, 12310-12316

SUPPLEMENTARY INFORMATION

Characterisation of class III biotin protein ligases; novel antifungal targets for human and agricultural pathogens

Louise M Sternicki[1], Stephanie Nguyen[1], Nicole R Pendini[1], Perdita Barran[2], Kamila Pacholarz[2], Tara L Pukala[3], Grant W Booker[1], Kate L Wegener[1, 4] and Steven W Polyak[1, 4, 5] *

[1] School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

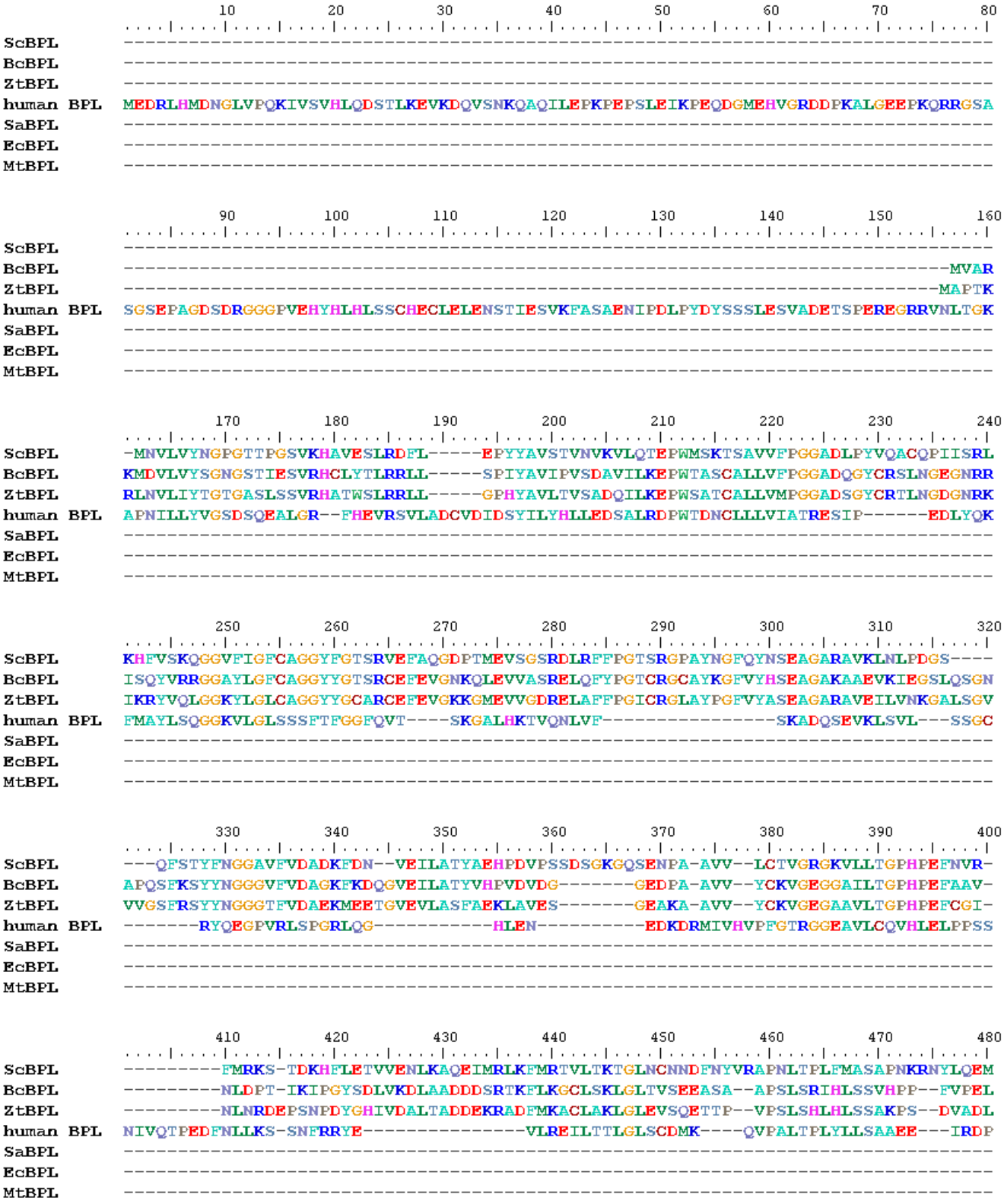
[2] Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, United Kingdom

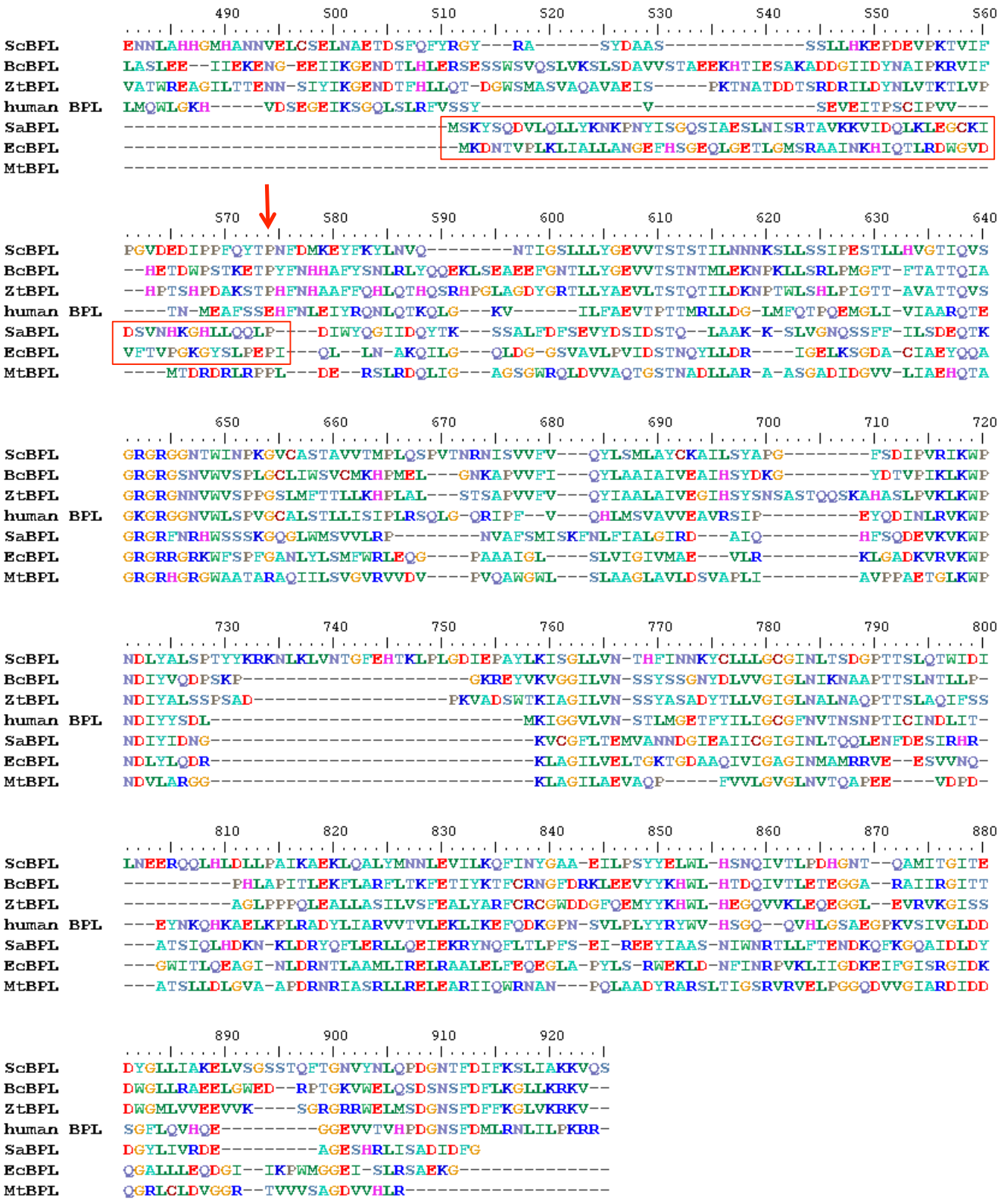
[3] School of Physical Sciences, The University of Adelaide, South Australia 5005, Australia

[4] Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

[5] Current Address: School of Pharmacy and Medical Sciences, University of South Australia, South Australia 5001, Australia

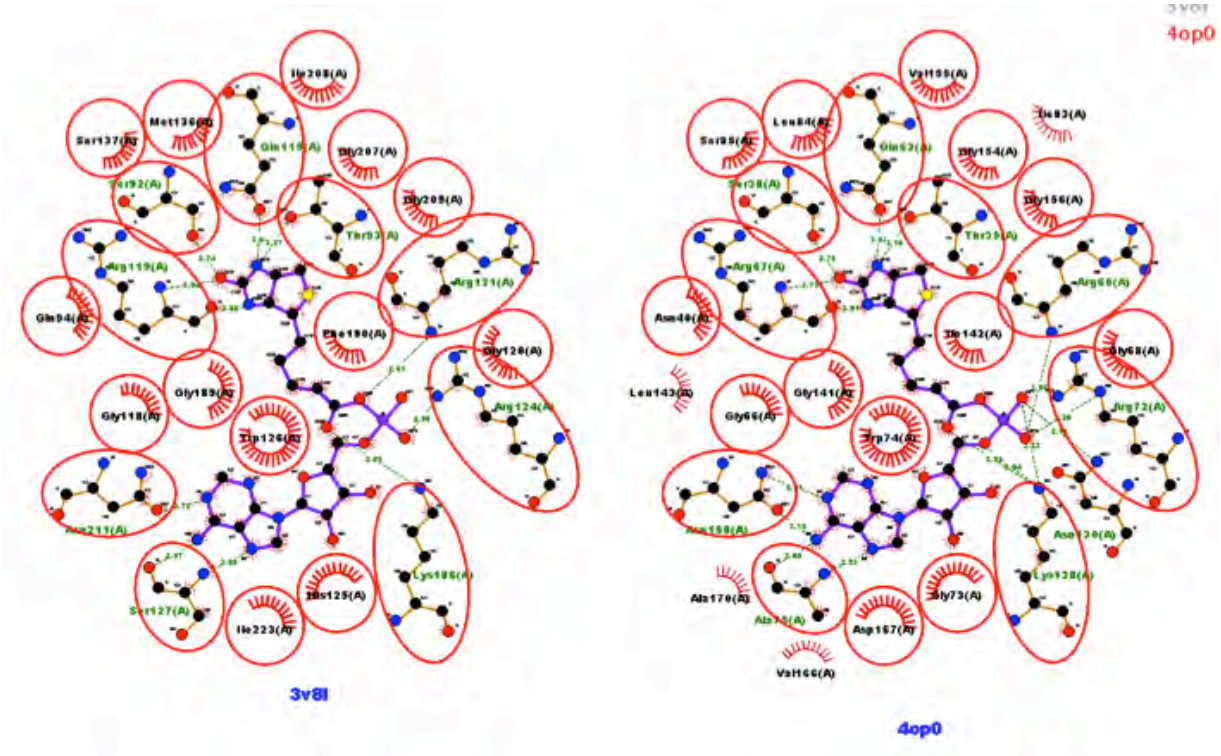
*To whom correspondence should be addressed.


ScBPL
BcBPL
ZtBPL
human BPL
SaBPL
EcBPL
MtBPL

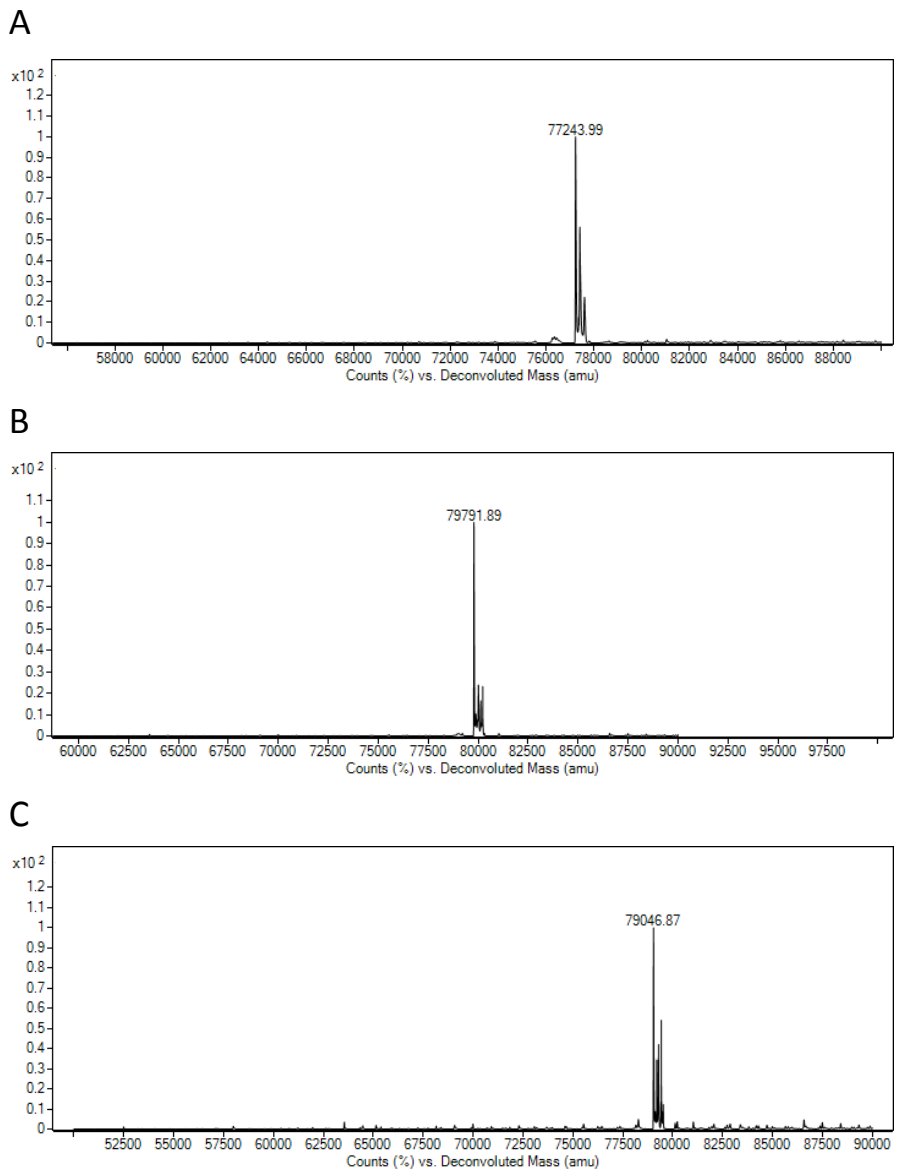


Supplementary Figure 1. Sequence alignment of the class III BPLs from *Saccharomyces cerevisiae* (*Sc*BPL), *Botrytis cinerea* (*Bc*BPL), *Zymoseptoria tritici* (*Zt*BPL) and humans. These are aligned with the bacterial BPLs from *Staphylococcus aureus* (*Sa*BPL), *Escherichia coli* (*Ec*BPL) and *Mycobacterium tuberculosis* (*Mt*BPL). Sequences are colour coded to identify identical residues. The red arrow represents the boundary between the N-terminal

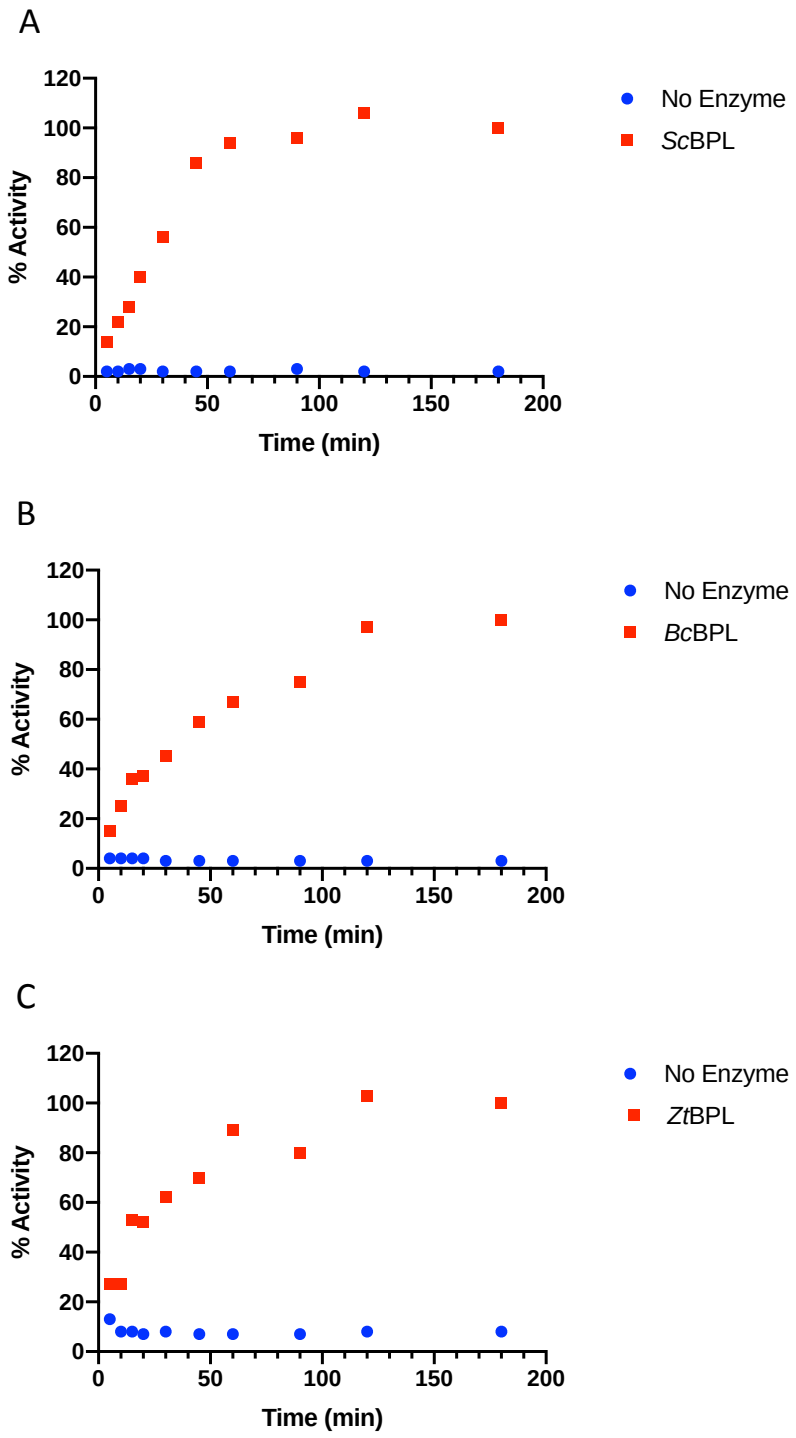
extension and the catalytic domain (and C-terminal cap) of class III BPLs. The red box identifies the DNA-binding domains of *Sa*BPL and *Ec*BPL that are not conserved amongst the other BPLs and, therefore, are to be disregarded from the alignment.



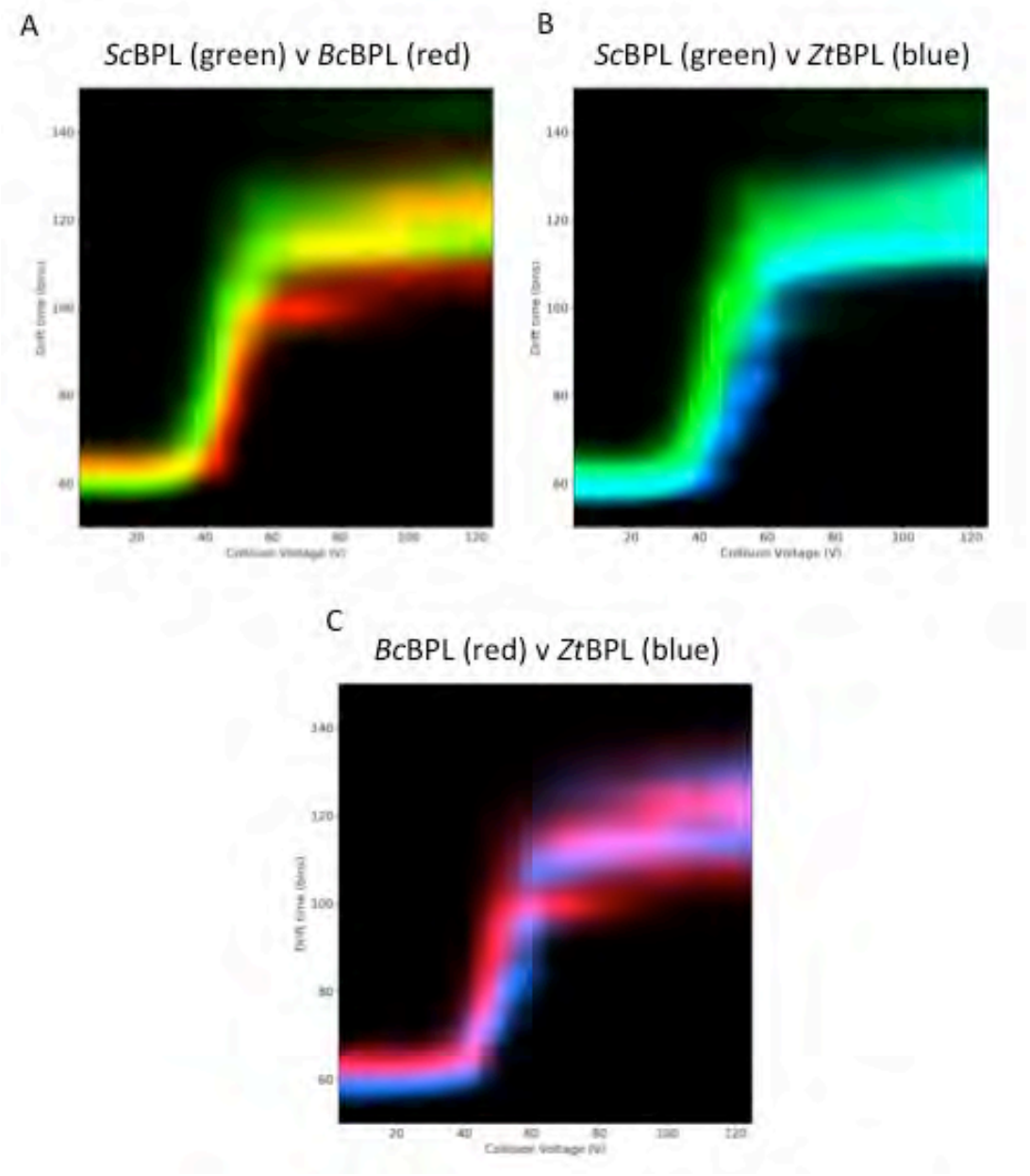
Supplementary Figure 2. Residues of the *S. aureus* and *M. tuberculosis* BPLs (PDB: 3v8l (1) and 4op0 (2) respectively) that interact with the reaction intermediate (purple) (identified using LigPlot+ (3,4)). Residues involved in hydrogen bond interactions are labelled green, and residues forming hydrophobic interactions are red. Circled residues are involved in ligand binding in both structures.



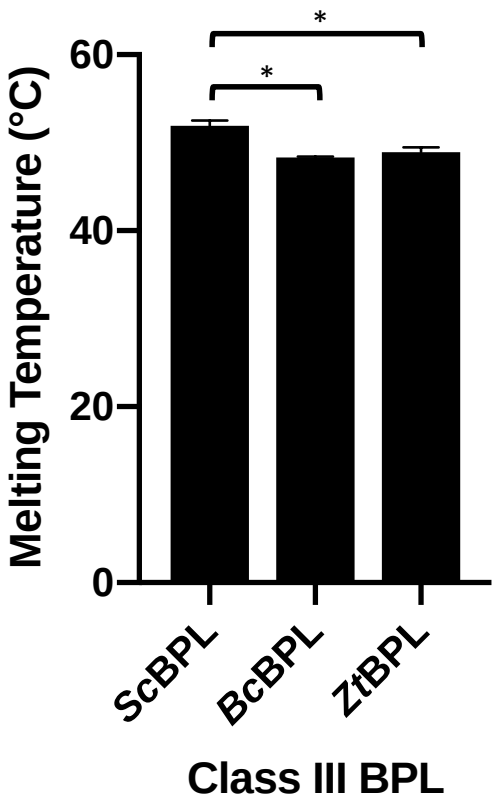
Supplementary Figure 3. Denatured MS of the class III BPLs from A) *S. cerevisiae,* B) *B. cinerea,* and C) *Z. tritici* to confirm the molecular mass of the recombinantly produced protein.



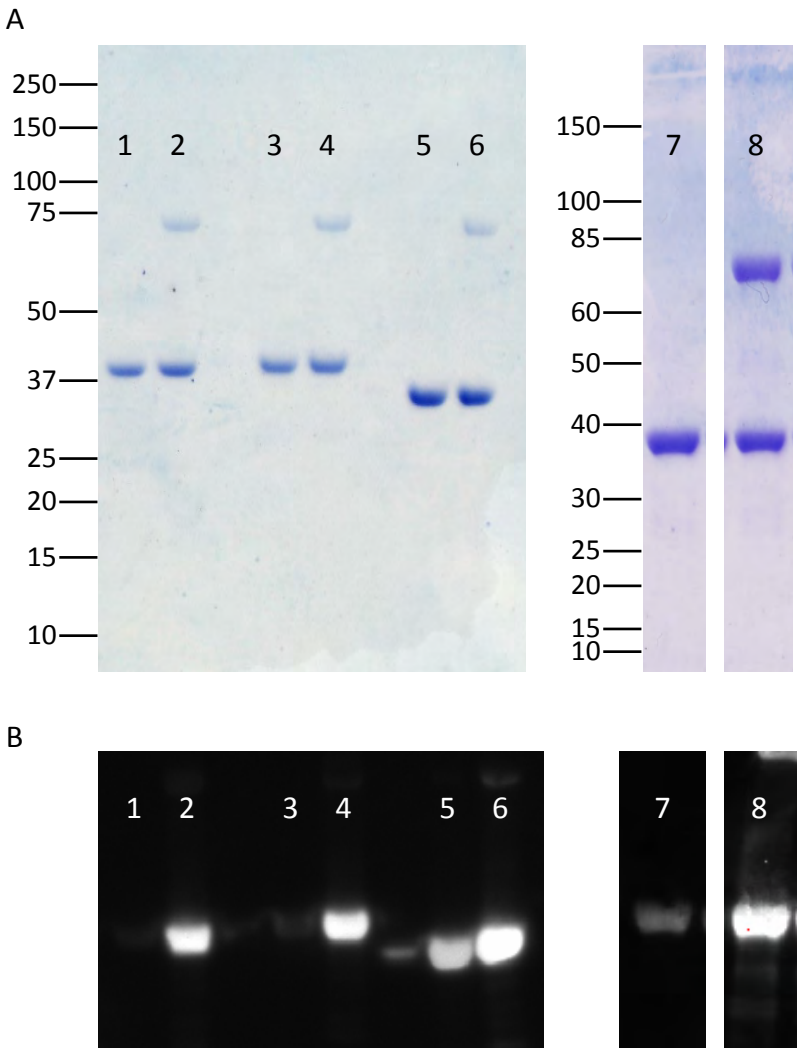
Supplementary Figure 4. The fungal class III enzymes from A) *S. cerevisiae*, B) *B. cinerea* and C) *Z. tritici* were active to biotinylate the biotin domain from the *S. cerevisiae* pyruvate carboxylase isoform 1 (GST-*Sc*PC104). The level of biotinylation was measured over time to identify the reaction time and enzyme concentration that allowed the reaction to be measured in the linear range. n=1.



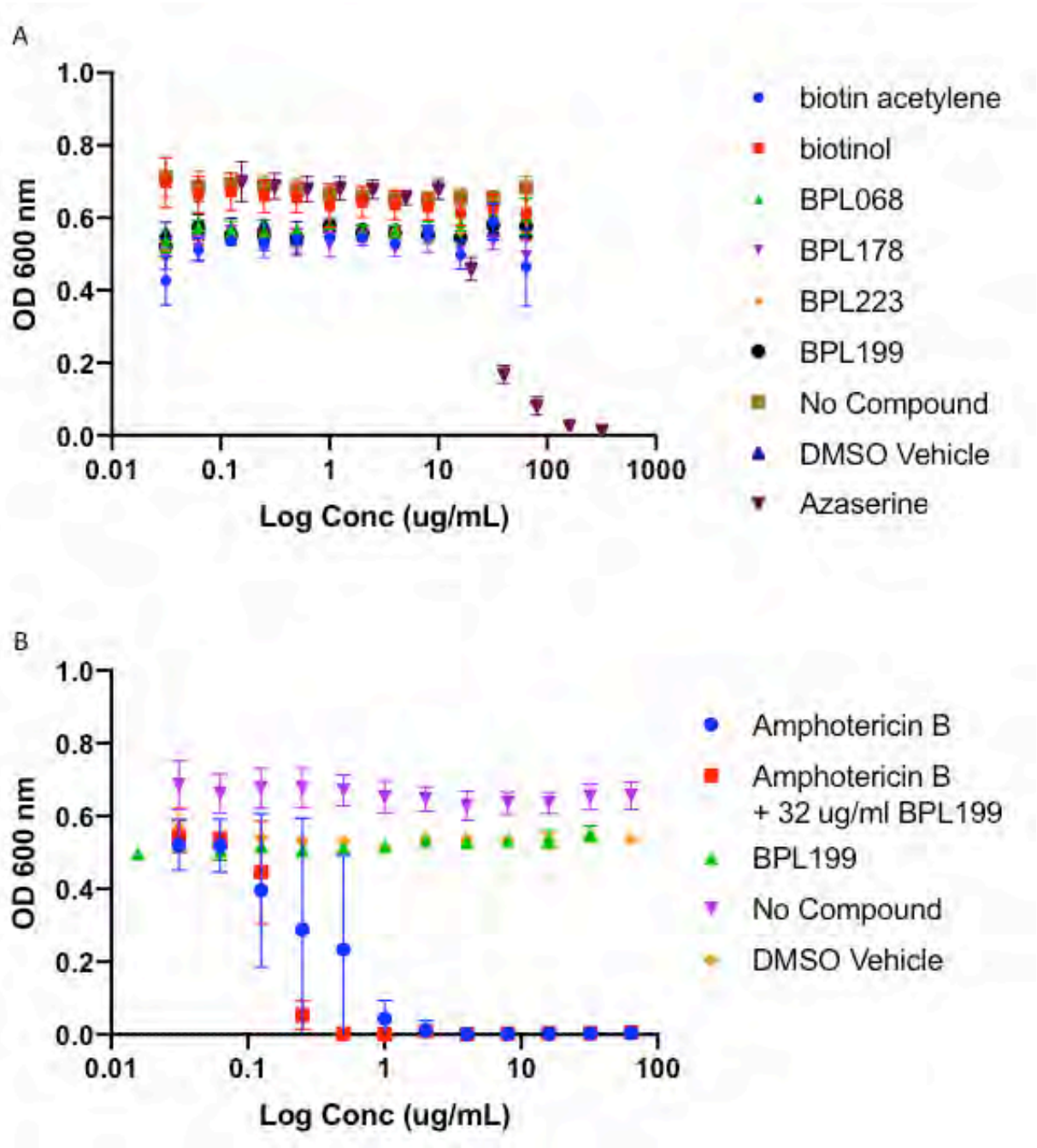
Supplementary Figure 5. Overlays of the CIU-MS denaturation curves for the 17+ charge state of the class III BPLs. Overlays compare A) *Sc*BPL (green) and *Bc*BPL (red), B) *Sc*BPL (green) and *Zt*BPL (blue), and C) *Bc*BPL (red) and *Zt*BPL (blue). Data were collected for increasing trap energies from 4 V to 200 V, but the figures were truncated at 125 V as there was no further unfolding past this voltage.



Supplementary Figure 6. *Sc*BPL is significantly more stable in in-solution thermal denaturation assays compared to *Bc*BPL and *Zt*BPL. The melting temperatures (the temperature where 50% of the protein is unfolded) are the average of 3 independent replicates, each with 3 technical replicates, plotted with the standard error of the mean. n=9.



Supplementary Figure 7. The biotin domains GST-*Sc*PC104 and GST-*Ca*PC115 were produced in their apo-state with no attached biotin, whilst GST-*Sa*PC90 and GST-*Bc*PC103 contained a proportion of biotinylated protein. B) Western blots were conducted using streptavidin-conjugated fluorophores to detect biotinylated biotin domains, together with the A) SDS-PAGE loading control. Lanes contain (1) purified GST-*Sc*PC104, (2) holo-GST-*Sc*PC104 control, (3) purified GST-*Ca*PC115, (4) holo-GST-*Ca*PC115 control, (5) purified GST-*Sa*PC90, (6) holo-GST-*Sa*PC90 control, (7) purified GST-*Bc*103, (8) holo-GST-*Bc*PC103 control.



Supplementary Figure 8. A) The BPL inhibitors have no antifungal activity against the *S. cerevisiae* w303 strain (n=1), B) even when in synergy with the positive control antifungal Amphotericin B (n=2). Azaserine and Amphotericin B are both known antifungal compounds that are utilised as positive controls. Results are the average of three technical replicates for each biological replicate with the background absorbance subtracted.

Supplementary Table 1. Sequences of the BPLs from *S. cerevisiae*, *B. cinerea* and *Z. tritici* used for recombinant protein production in this study.

Species of BPL	**Protein sequence utilised in this study**	**Reference of sequence identification**
Saccharomyces cerevisiae	MNVLVYNGPGTTPGSVKHAVESLRDFLEPYYAVSTVNVKVLQTEPWMSKTSAVVFPGGADLPYVQACQPIISRLKH FVSKQGGVFIGFCAGGYFGTSRVEFAQGDPTMEVSGSRDLRFFPGTSRGPAYNGFQYNSEAGARAVKLNLPDGSQ FSTYFNGGAVFVDADKFDNVEILATYAEHPDVPSSDSGKGQSENPAAVVLCTVGRGKVLLTGPHPEFNVRFMRKST DKHFLETVVENLKAQEIMRLKFMRTVLTKTGLNCNNDFNYVRAPNLTPLFMASAPNKRNYLQEMENNLAHHGMHAN NVELCSELNAETDSFQFYRGYRASYDAASSSLLHKEPDEVPKTVIFPGVDEDIPPFQYTPNFDMKEYFKYLNVQNTIG SLLLYGEVVTSTSTILNNNKSLLSSIPESTLLHVGTIQVSGRGRGGNTWINPKGVCASTAVVTMPLQSPVTNRNISVVF VQYLSMLAYCKAILSYAPGFSDIPVRIKWPNDLYALSPTYYKRKNLKLVNTGFEHTKLPLGDIEPAYLKISGLLVNTHFIN NKYCLLLGCGINLTSDGPTTSLQTWIDILNEERQQLHLDLLPAIKAEKLQALYMNNLEVILKQFINYGAAEILPSYYELWL HSNQIVTLPDHGNTQAMITGITEDYGLLIAKELVSGSSTQFTGNVYNLQPDGNTFDIFKSLIAKKVQSHHHHHH	(Polyak *et al.* 1999 JBC) (5)
Botrytis cinerea	MHHHHHHENLYFQGSMVARKMDVLVYSGNGSTIESVRHCLYTLRRLLSPIYAVIPVSDAVILKEPWTASCALLVFPGGA DQGYCRSLNGEGNRRISQYVRRGGAYLGFCAGGYYGTSRCEFEVGNKQLEVVASRELQFYPGTCRGCAYKGFVYHS EAGAKAAEVKIEGSLQSGNAPQSFKSYYNGGGVFVDAGKFKDQGVEILATYVHPVDVDGGEDPAAVVYCKVGEGGAI LTGPHPEFAAVNLDPTIKIPGYSDLVKDLAADDDSRTKFLKGCLSKLGLTVSEEASAAPSLSRIHLSSVHPPFVPELLASL EEIIEKENGEEIIKGENDTLHLERSESSWSVQSLVKSLSDAVVSTAEEKHTIESAKADDGIIDYNAIPKRVIFHETDWPSTK ETPYFNHHAFYSNLRLYQQEKLSEAEEFGNTLLYGEVVTSTNTMLEKNPKLLSRLPMGFTFTATTQIAGRGRGSNVWV SPLGCLIWSVCMKHPMELGNKAPVVFIQYLAAIAIVEAIHSYDKGYDTVPIKLKWPNDIYVQDPSKPGKREYVKVGGIL VNSSYSSGNYDLVVGIGLNIKNAAPTTSLNTLLPPHLAPITLEKFLARFLTKFETIYKTFCRNGFDRKLEEVYYKHWLH TDQIVTLETEGGARAIIRGITTDWGLLRAEELGWEDRPTGKVWELQSDSNSFDFLKGLLKRKVGSHHHHHHGSAGS AKKKGSAGSAHHHHHH	This study
Zymoseptoria tritici	MHHHHHHENLYFQGSMAPTKRLNVLIYTGTGASLSSVRHATWSLRRLLGPHYAVLTVSADQILKEPWSATCALLVM PGGADSGYCRTLNGDGNRKIKRYVQLGGKYLGLCAGGYYGCARCEFEVGKKGMEVVGDRELAFFPGICRGLAYP GFVYASEAGARAVEILVNKGALSGVVVGSFRSYYNGGGTFVDAEKMEETGVEVLASFAEKLAVESGEAKAAVVYC KVGEGAAVLTGPHPEFCGINLNRDEPSNPDYGHIVDALTADDEKRADFMKACLAKLGLEVSQETTPVPSLSHLHLS SAKPSDVADLVATWREAGILTTENNSIYIKGENDTFHLLQTDGWSMASVAQAVAEISPKTNATDDTSRDRILDYNLV TKTLVPHPTSHPDAKSTPHFNHAAFFQHLQTHQSRHPGLAGDYGRTLLYAEVLTSTQTILDKNPTWLSHLPIGTTAV ATTQVSGRGRGNNVWVSPPGSLMFTTLLKHPLALSTSAPVVFVQYIAALAIVEGIHSYSNSASTQQSKAHASLPVKL KWPNDIYALSSPSADPKVADSWTKIAGILVNSSYASADYTLLVGIGLNALNAQPTTSLAQIFSSAGLPPPQLEALLASIL VSFEALYARFCRCGWDDGFQEMYYKHWLHEGQVVKLEQEGGLEVRVKGISSDWGMLVVEEVVKSGRGRRWELM SDGNSFDFFKGLVKRKVGSHHHHHHGSAGSAKKKGSAGSAHHHHHH	This study

Supplementary Table 2. The hydrogen bond interactions the *M. tuberculosis* BPL (*Mt*BPL) (PDB 4op0) (2) and *S. aureus* BPL (*Sa*BPL) (PDB 3v8l) (1) make with the reaction intermediate, as determined by LigPlot+ (3,4). The part of the reaction intermediate each residue interacts with is annotated.

	3v8l	**4op0**	
	***Sa*BPL**	***Mt*BPL**	
side chain	Ser92	Ser38	Biotin
backbone	Thr93	Thr39	
backbone	Gln115	Gln63	
side chain	Arg119	Arg67	
backbone	Arg121	Arg69	Phosphodiester / Linker
side chain	Arg124	Arg72	
side chain	Lys186	Lys138	
backbone	Ser127	Ala75	ATP (adenine)
side chain	Asn211	Asn158	

Supplementary Table 3. The hydrophobic interactions between the reaction intermediate and both *M. tuberculosis* BPL (*Mt*BPL) (PDB 4op0) (2) and *S. aureus* BPL (*Sa*BPL) (PDB 3v8l) (1) as determined by LigPlot+ (3,4).

3v8l	**4op0**
***Sa*BPL**	***Mt*BPL**
Gln94	Asn40
Gly118	Gly66
Gly120	Gly68
His125	Gly73
Trp126	Trp74
	Ile83
Met136	Leu84
Ser137	Ser85
Gly189	Gly141
Phe190	Ile142
	Leu143
Gly207	Gly154
Ile208	Val155
Gly209	Gly156
	Val166
Ile223	Asp167
	Ala170

Supplementary Table 4. Mean, standard deviation (SD) and full width half maximum (FWHM) calculated from the IM-MS CCS distribution data for the 17+ and 18+ charge states of the three fungal BPLs, *Sc*BPL, *Bc*BPL and *Zt*BPL.

		***Sc*BPL**	***Bc*BPL**	***Zt*BPL**
18+	MEAN (nm^2)	52.25	52.06	51.85
	SD (nm^2)	0.87	1.13	1.09
	FWHM (nm^2)	1.34	2.35	2.04
17+	MEAN (nm^2)	50.42	50.83	50.28
	SD (nm^2)	0.87	1.25	1.12
	FWHM (nm^2)	1.84	2.75	2.48

Supplementary Table 5. The melting temperatures of the fungal class III BPLs measured in solution by thermal denaturation assays. The melting temperatures are the average of 3 independent replicates, each with 3 technical replicates, plus/minus the standard error of the mean. n=9.

Class III BPL	**T_M ± SEM (°C)**
*Sc*BPL	51.9 ± 0.4
*Bc*BPL	48.3 ± 0.1
*Zt*BPL	48.9 ± 0.3

References:

1. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterisation of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci* **22**, 762-773
2. Ma, Q., Akhter, Y., Wilmanns, M., and Ehebauer, M. T. (2014) Active site conformational changes upon reaction intermediate biotinyl-5′ -AMP binding in biotin protein ligase from *Mycobacterium tuberculosis*. *Protein Sci* **23**, 932-939
3. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-134
4. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786
5. Polyak, S. W., Chapman-Smith, A., Brautigan, P. J., and Wallace, J. C. (1999) Biotin Protein Ligase from *Saccharomyces cerevisiae*. The N-terminal domain is required for complete activity. *J Biol Chem* **274**, 32847-32854

CHAPTER 5:

STRUCTURAL INVESTIGATION OF THE CLASS III *Saccharomyces cerevisiae* BPL

CHAPTER OUTLINE

This manuscript style chapter further investigates the structure of a class III BPL, using the BPL from the prototypical model organism *S. cerevisiae.* Native nanoelectrospray ionisation-mass spectrometry (nESI-MS), which was optimised in the previous chapter, was employed here. nESI-MS allowed the use of ion mobility and collision-induced unfolding-MS experiments, which together with the non-MS techniques of circular dichroism and thermal denaturation assays, permitted the investigation of the structural differences that occur upon the binding of the ligands biotin and MgATP, along with the influence of these ligands on the stability of the enzyme. Homology modelling and hydrogen-deuterium exchange MS were utilised to reveal novel insights into the structure of the BPL, particularly the structure of the domain located in the N-terminal extension of class III BPLs. Overall, this chapter aims to investigate the structure of class III BPLs in order to 1) better understand this unique class of BPLs and 2) provide structural insights into fungal BPLs to assist with the development of antifungal agents.

STATEMENT OF AUTHORSHIP

Statement of Authorship

Title of Paper	Structural investigations of the N-terminal domain of the class III *Saccharomyces cerevisiae* Biotin Protein Ligase
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Louise M Sternicki
Contribution to the Paper	Helped conceive the project and designed the experiments. Produced recombinant proteins and completed all experiements. Data analysis for all experiments. Manuscript preparation.
Overall percentage (%)	85%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	[redacted] Date 6/3/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Tara L Pukala
Contribution to the Paper	Advised on mass spectrometry experiments and assisted in collecting data. Assisted in analysing data and helped prepare the manuscript.
Signature	[redacted] Date 6/3/19

Name of Co-Author	Kamila Pacholarz
Contribution to the Paper	Advised on mass spectrometry experiments and assisted with data collection.
Signature	[redacted] Date 7/10/2019

Name of Co-Author	Perdita Barran		
Contribution to the Paper	Advised on mass spectrometry experiments.		
Signature		Date	24/7/19

Name of Co-Author	Grant W Booker		
Contribution to the Paper	Provided advice on the project design, data analysis, and assited with manuscript preparation.		
Signature		Date	6/3/2019

Name of Co-Author	Kate L Wegener		
Contribution to the Paper	Assisted with project and experimental design, data analysis, and preparation of the manuscript.		
Signature		Date	6/3/19

Name of Co-Author	Steven W Polyak		
Contribution to the Paper	Conceived the project, assisted with project and experimental design, and helped prepare the manuscript.		
Signature		Date	6/3/19

MANUSCRIPT

Structural investigations of the N-terminal domain of the class III *Saccharomyces cerevisiae* Biotin Protein Ligase

Louise M Sternicki[1]*, Tara L Pukala[2], Kamila Pacholarz[3], Perdita Barran[3], Grant W Booker[1], Steven W Polyak[1, 4, 5 +] and Kate L Wegener[1, 4 *+]

[1] School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

[2] School of Physical Sciences, The University of Adelaide, South Australia 5005, Australia

[3] Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, United Kingdom

[4] Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

[5] Current address: School of Pharmacy and Medical Sciences, University of South Australia, 5001, Australia

* To whom correspondence should be addressed.

[+] Equal senior authors

Abstract

Biotin (vitamin B7) is an essential co-factor for the activity of particular metabolic enzymes from all forms of life. Biotin protein ligase is the enzyme responsible for the covalent attachment of biotin onto the biotin domain of these biotin-dependent enzymes. Certain eukaryotic BPLs are distinct from those of bacteria, archaea and plants, due to the presence of a unique, large N-terminal extension in addition to the conserved catalytic domain and C-terminal cap. No structures of a eukaryotic BPL have been solved, however, previous functional studies have revealed the N-terminal extension interacts with the biotinylation substrate. Here we have utilised mass spectrometry (MS) techniques to investigate the structure of the yeast *Saccharomyces cerevisiae* BPL (*Sc*BPL), revealing the conformational differences between apo- and ligand bound-*Sc*BPL. Whilst lower resolution techniques revealed no gross structural changes occur with ligand binding, holo-*Sc*BPL was shown to have a more compact and less dynamic conformation that resulted in an increased stability. The higher-resolution technique of hydrogen-deuterium exchange (HDX) coupled to MS was utilised together with homology modelling to further inform on the structure of *Sc*BPL. A folded domain was identified in the N-terminal extension that evidence suggests has the fold of a glutamine amidotransferase. A mostly solvent protected stretch of 160-residues is suggested to form a folded, partially buried linker that stabilises the interface between the N-terminal extension and the catalytic domain. Further HDX analysis identified localised conformational changes in both the active site and N-terminal domain of *Sc*BPL that occur concomitantly with ligand binding, as well as potential interaction interfaces between the two domains and/or the linker. Together these data provide novel insights into the unique structure of a class III BPL and how ligands influence this structure for catalysis of protein biotinylation.

Introduction

Biotin is essential for all forms of life as it is an indispensible covalent cofactor for key metabolic enzymes, where it acts as a transient carrier of carboxyl groups during metabolic reactions (1). In humans, impairment in the activity of these enzymes results in Multiple Carboxylase Deficiency (MCD), a rare but serious disease causing a range of severe symptoms in neonates or juveniles, ranging from complications and delays in development, to coma and death (2-4).

Biotin protein ligase (BPL) is the enzyme responsible for the post-translational covalent attachment of biotin to biotin-dependent enzymes, attaching it to a single conserved lysine residue in the biotin domain of these enzymes (5). BPL carries out this modification via a two-step mechanism. The first step involves the binding of biotin and MgATP in adjacent pockets in the active site where they ligate to form the intermediate biotinyl-5′-AMP (6). BPL then interacts with the biotin domain of the biotin-dependent enzyme, correctly positioning the conserved target lysine residue adjacent to the BPL active site such that the biotinyl moiety can be transferred and AMP released (7). Some BPLs require biotin to bind prior to ATP, as the binding of biotin induces conformational changes in the active site to form the ATP binding site (8-12). This includes the ordering of a previously disordered loop (termed the biotin-binding loop), which re-positions a key tryptophan residue to accommodate pi-pi stacking interactions with the purine rings of ATP (11,13). The ordering of additional loops in the active site upon ligand binding further stabilises the binding of ligands and protects the labile reaction intermediate from hydrolysis and dissociation (11,14).

BPLs across all kingdoms of life can be divided into three structural classes. The class I and II enzymes are the simplest enzymes and are mainly found in archaea, bacteria and plants. Class I enzymes contain the catalytic domain and C-terminal cap, which together constitute the minimal structural requirements for protein biotinylation (6,9,15,16). In addition to these modules, class II BPLs also contain an N-terminal DNA binding domain (11,13,17), which allows transcriptional regulation of biotin-related genes, such as those involved in biotin transport and synthesis, in response to intracellular biotin concentrations (reviewed (18)). In contrast, the class III enzymes from fungi, mammals and insects contain a much larger N-terminal extension in addition to the catalytic domain and C-terminal cap. This extension has no homology with the class II N-terminal domain and no predicted DNA binding capability. While the class I and II enzymes have been thoroughly characterised, with multiple structures solved from each class, the class III enzymes have been less well studied. There is no atomic

resolution structure of a class III enzyme, and no definitive function has been assigned to the large N-terminal extension. There have been many attempts to solve the structure of a class III BPL using X-ray crystallography, but to date these efforts have been unsuccessful.

Previous structural studies on class III BPLs have concentrated on the yeast and human enzymes. Limited proteolysis experiments and functional assays with truncation mutants suggested the presence of a structured domain in the N-terminal extension of both these enzymes that was essential for activity (12,19-21), located between residues 1-220 in *Sc*BPL (19) and 160-300 in the human BPL (12,20). The N-terminal extension has been hypothesised to promote BPL activity by interacting with the biotin domain to form a stable, enduring interaction for biotinylation to occur. Evidence of this interaction has come from yeast-two hybrid (21) and surface plasmon resonance (SPR) (12) experiments on *Sc*BPL and human BPL respectively. Nuclear magnetic resonance (NMR) binding studies showed that the 160 residues preceding the predicted N-terminal structured domain in human BPL also interact with the biotin domain (22). In addition to stabilising the substrate interaction, it has also been proposed that the N-terminal extension promotes 'substrate verification' activity, whereby only appropriate biotin domains are selected for modification (23). Several studies have demonstrated class III BPLs are selective towards the biotin domain substrates they natively biotinylate (reviewed (23)) in preference to biotin domains originating from different species (20,24,25) (Chapter 4), or other biotin-dependent enzymes within a species (26). Alternate isoforms of human BPL that begin translation initiation at different N-terminal residues (M1 versus M57) had different affinities for a biotin domain substrate (27). This provided further evidence of a role for the N-terminal extension in substrate recognition and selectivity. Despite this, no specific structural features in the conserved folds of biotin domains (28) have been identified that confer selective recognition by class III BPLs. Only one atomic resolution structure of a BPL in complex with its biotin domain substrate has been solved (7). However, this is a class I BPL from the thermophilic bacteria *Pyrococcus horikoshii*, and hence does not contain an N-terminal extension. Whilst the interaction of the catalytic domain with the biotin domain is predicted to be conserved amongst different BPL classes, the structure, interactions, and positioning of the N-terminal domain in this interaction remains unknown. Likewise, none of the current structural data have helped elucidate the mechanism class III BPLs use for selecting appropriate substrates.

The biotinylation mechanism employed by eukaryotic BPLs is proposed to include conformational changes to the enzyme concomitant with ligand binding. This has been

hypothesised based on limited proteolysis experiments of the *S. cerevisiae* class III BPL, for which a reduction in proteolysis was observed upon the addition of the ligands biotin and MgATP (19). These low-resolution data implied ligand-induced conformational changes alter the structure of the protease sensitive region (residues 220-260) such that it is less accessible to proteases (19). Further structural and functional investigation of ligand-induced rearrangements in *Sc*BPL is required to map the precise locations of these and other potential conformational changes, how these regions change structurally, and to understand how these contribute to the biotinylation reaction of class III BPLs.

Uncovering the mechanism of class III BPLs will be crucial for understanding the clinical manifestations of MCD. Of particular interest are MCD missense mutations that do not respond to supplemental biotin therapy (29-31) and do not cause disease by increasing the K_M for biotin (32-34). Many of the relevant mutations for these MCD variants map to the unique N-terminal extension of the human BPL and have been suggested to cause disease by reducing the affinity of the enzyme for the biotin domain substrate through increasing the dissociation rate, thereby, preventing an enduring interaction of the complex (12). Hence, structural investigation of class III BPLs and understanding the involvement of the N-terminal extension in their mechanism of action is crucial for understanding the structural basis of these mutations.

Here we reveal ligand binding induced structural changes in the class III BPL from the model organism *S. cerevisiae*. Low-resolution structural techniques, such as circular dichroism (CD) and ion mobility-mass spectrometry (IM-MS), did not identify any gross conformational changes to the structure of *Sc*BPL upon ligand binding. Stability of the enzyme did increase in the presence of ligands, with IM-MS suggesting a reduction in the dynamic movement of the enzyme in solvent and/or compaction of the enzyme's structure when in complex with biotinyl-5′-AMP. The higher-resolution method of hydrogen-deuterium exchange (HDX), detected via liquid chromatography-mass spectrometry (LC-MS), revealed smaller, localised conformational changes within *Sc*BPL concurrent with ligand binding. The HDX results together with homology modelling also allowed us to propose a new structural model of the eukaryotic class III BPL from *S. cerevisiase* in the absence of an empirically determined atomic resolution structure and map the locations of these localised structural changes. These data provide a basis for understanding the structure of a class III BPL, which will enable further investigation towards understanding differences in the function of this complex class of BPLs compared to the well-studied class I/II enzymes.

Methods

Recombinant production of Apo-*S. cerevisiae* BPL

Apo-*Sc*BPL expression and purification was performed similarly to that described in (35). A 10 mL starter culture of w303 *S. cerevisiae* harbouring the pVT100u expression vector for *Sc*BPL-H_6 was grown in yeast minimal media (8 g/L Yeast Nitrogen Base without amino acids and ammonium sulphate (Difco), 11 g/L Cassamino Acids (Difco), 1.48 mM adenine, 0.01% tryptophan, 0.01% leucine, 2% glucose, 41 µM *d*-biotin) at 30 °C for 24 to 48 hours with agitation at 170 rpm. A 1:100 dilution of this starter culture was used to inoculate a 100 mL sub-culture of yeast minimal media that was grown at 30 °C for 24 hours with agitation at 170 rpm. Cells were pelleted by centrifugation at 3300 x *g* for 5 min before re-suspension in 10 mL yeast minimal media to inoculate 500 mL of fresh yeast minimal media, which was then grown at 30 °C for 48 hours with agitation at 170 rpm. Cells were harvested by centrifugation at 1900 x *g* and 4 °C for 10 minutes before freezing the pellets at -80 °C. The cell pellets were washed with 1x PBS and re-harvested by repeating the centrifugation before freezing the pellets at -80 °C. Cells were then washed in wash buffer 1 (10 mM Tris pH 7.5, 0.65 M sorbitol, 0.1 mM EDTA, 0.1 mM DTT) three times, before being resuspended in wash buffer 2 (10 mM Tris pH 7.5, 0.65 M sorbitol, 0.1 mM EDTA, 1 mM DTT) supplemented with 1000 units of lyticase. This cell suspension was incubated for 1.5 hours at 30 °C with agitation at 150 rpm, before re-harvesting by centrifugation at 1900 x *g* for 15 minutes.

Cell lysates were prepared by chemical lysis using 5 mL of Yeast Buster ® (Novagen) per gram of cell pellet together with 1 mM Tris(hydroxypropyl)phosphine, 1 mM PMSF and cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche) at room temperature with gentle agitation for 1 hour. Lysate was then clarified by centrifugation at 1900 x *g* for 15 minutes and 0.45 µm filtration. Clarified lysate was then treated as previously outlined (36) to allow the purification of apo-*Sc*BPL. This lysate was applied to a 5 mL HiPrep™ IMAC Fast Flow column (GE Healthcare) pre-charged with nickel sulphate and equilibrated with NiNTA Wash Buffer (300 mM NaCl, 50 mM NaH_2PO_4, 20 mM imidazole) for purification of the hexa-histidine tagged *Sc*BPL. Following application of the lysate, the column was washed with NiNTA Wash Buffer, followed by 50 mM imidazole using 10% NiNTA Elution Buffer (300 mM NaCl, 50 mM NaH_2PO_4, 500 mM imidazole) in NiNTA Wash Buffer. Finally, *Sc*BPL-H_6 was eluted with 200 mM imidazole, using 40% NiNTA Elution Buffer in NiNTA Wash Buffer. IMAC-purified *Sc*BPL was buffer exchanged into BPL Storage Buffer (50 mM Tris pH 8.0, 5% glycerol, 1 mM DTT, 0.5 mM EDTA pH 8.0) via dialysis overnight at 4°C.

*Sc*BPL was resolved from residual contaminating proteins by anion exchange chromatography. The dialysed *Sc*BPL was loaded onto an 8 mL ENRICH Q-Sepharose column (Bio-Rad) that was pre-equilibrated with BPL Storage Buffer. Following column washing with BPL Storage Buffer, *Sc*BPL was eluted with a 0-20% gradient of BPL Storage Buffer with 1 M NaCl. The fractions containing *Sc*BPL were pooled and buffer exchanged via dialysis overnight at 4 °C into BPL Storage Buffer to remove excess salt. *Sc*BPL was concentrated using Vivaspin® 20 30,000 MWCO PES spin concentrators (Sartorius) and stored at -80 °C. SDS-PAGE was utilised to assess the purity of *Sc*BPL.

Streptavidin blot to measure the apo-state of *Sc*BPL

The amount of biotin or reaction intermediate bound to *Sc*BPL following expression and purification was determined using the Streptavidin blot method previously reported (36,37). The reactions completed prior to the blot contained 3 µM or 15 µM of the *S. cerevisiae* biotin domain substrate from pyruvate carboxylase isoform 1 (GST-*Sc*PC104), 3 µM of apo-*Sc*BPL, and 3 µM or 15 µM biotin.

Native nano-electrospray ionisation ion mobility-mass spectrometry (native nESI IM-MS)

*Sc*BPL was buffer exchanged into 1 M ammonium acetate pH 6.9, then 100 mM ammonium acetate pH 6.9 using three sequential Micro Bio-Spin™ P-6 Gel Columns (Bio-Rad). To produce holo-*Sc*BPL, *Sc*BPL was incubated with 500 µM biotin, 1 mM $MgCl_2$, and 1 mM ATP on ice for at least 1 hour prior to buffer exchange. β-lactoglobulin, avidin, albumin, concanavilin A and alcohol dehydrogenase were used as collision cross section (CCS) calibrants (38). These proteins were buffered exchanged in the same manner as *Sc*BPL. All proteins were quantified using a NanoDrop (Thermo Scientific) with molecular weights and extinction coefficients predicted from the protein sequences via the ProtParam Tool on the ExPASy Bioinformatics Research Portal. Proteins were diluted to 10 µM in 100 mM ammonium acetate for MS analysis.

MS analysis was completed using a Synapt G2 S High Definition Mass Spectrometer (HDMS) (Waters Corporation) with samples introduced in the positive ion mode from a nano-electrospray ionisation (nESI) source. The positive potential was applied to the sample along a platinum wire inserted into a glass capillary that was prepared in-house. Parameters were optimised to maintain non-covalent interactions and native protein conformations, and

included capillary voltage, 1.0 – 1.6 kV; sampling cone, 50 V; source temperature, 40°C; trap collision energy, 4 V; transfer collision energy, 2 V; trap gas flow, 2 mL/min; and backing pressure 3.67 mbar. The specific ion mobility parameters included IM cell wave height, 35 V; IM cell wave velocity, 400 m/sec; transfer t-wave height, 0 V; transfer t-wave velocity, 191 m/sec. Data analysis was performed using MassLynx (Waters) with manual peak finding. Drift times were extracted using MassLynx V4.1 and DriftScope^{TM} 2.8 (Waters Corporation) for calculation of CCS values according to published calculations (39) using Microsoft Excel.

Collision induced unfolding-mass spectrometry (CIU-MS)

Samples were prepared and introduced into a Synapt G2 S HDMS (Waters) as for IM-MS. In-house software (ORIGAMI^{MS}) (40) was used to automatically control the trap collision energy, increasing this parameter in increments of 2 V over a range of 4 V to 200 V. For each voltage, 5 scans were collected with each scan 5 seconds. Data were collected in the MSMS mode selecting the *m/z* ion of the charge state of interest. Collected data were analysed via ORIGAMI^{ANALYSE} (40).

Circular dichroism (CD)

Apo-*Sc*BPL was buffer exchanged into 10 mM ammonium acetate pH 6.95 using Amicon Ultra-0.5 MWCO 10,000 centrifugal filter units (Merck). Holo-*Sc*BPL was formed by buffer exchanging apo-*Sc*BPL into 10 mM ammonium acetate pH 6.95 with 50 µM biotin, 100 µM $MgCl_2$ and 100 µM ATP. Protein concentrations were determined using a NanoDrop (ThermoFisher Scientific) (as for native IM-MS) and proteins diluted to 2.5 µM. Far UV CD measurements were completed on a Jasco CD J-815 spectropolarimeter at room temperature. Spectra were collected from 195 to 260 nm with 0.2 nm increments utilising a 10 mm path length cell. Results are the average of 5 scans corrected with buffer blanks and smoothed using Savitzky-Golay convolution. Spectra were exported from Jasco Spectral Analysis software for analysis with Microsoft Excel and GraphPad Prism.

Solution thermal denaturation assays to measure melting temperatures

Thermal denaturation assays were completed according to Boivin *et al.* (41) and Reinhard *et al.* (42). Holo-*Sc*BPL was formed by incubating 10 µM apo-*Sc*BPL with 50 µM biotin, 100 µM $MgCl_2$ and 100 µM ATP on ice for 30 minutes. Apo- and holo-*Sc*BPL were diluted 10-fold in 100 mM ammonium acetate and 5x SYPRO® Orange protein gel stain (ThermoFisher Scientific), giving final concentrations of 1 µM apo- or holo-*Sc*BPL and 5 µM biotin, 10 µM $MgCl_2$ and 10 µM ATP in the holo-*Sc*BPL samples. These reactions were plated into an

Applied Biosystems MicroAmp™ Fast Optical 96-well reaction plate (ThermoFisher Scientific) with a final well volume of 25 µL. Samples were plated in triplicate, together with an apo- and holo-*Sc*BPL only control and a dye only control. Plates were centrifuged at low speed (approximately 1000 x *g*) for 1 minute before sealing with an Applied Biosystems MicroAmp™ Optical Adhesive Film plate seal (ThermoFisher Scientific). A Step One Plus qPCR instrument (Applied Bioscience) was used to perform the melting curve, measuring the fluorescence using the ROX reporter channel every 0.3 °C, from 4 °C to 90 °C. Data were exported into Microsoft Excel before using GraphPad Prism to fit a Boltzman-Sigmoidal equation to the melting curve. This fit determined the melting point (T_M), which is the inflection point (V_{50}) of the melting curve. Each replicate was independently fit, and then the triplicate T_M values were averaged.

Hydrogen deuterium exchange mass spectrometry (HDX-MS)

*Sc*BPL was prepared via buffer exchange as for native IM-MS and diluted to 34 µM, however ligands were not added prior to buffer exchange. Instead, 200 µM biotin, 500 µM $MgCl_2$ and 500 µM ATP were added after buffer exchange to the holo-*Sc*BPL sample so that excess ligand was present during the labelling reaction. Deuterium labelling and quenching reactions were performed using the automatic CTC PAL sample manager (LEAP Technologies). *Sc*BPL was diluted approximately 15-fold in 10 mM potassium phosphate in 99.99% deuterium oxide pH 6.6 (pD 7.0) and incubated at 20 °C for varying time points of 30 seconds, 1 minute, 10 minutes, 30 minutes, 1 hour and 4 hours. A time point of 0 was included by diluting *Sc*BPL in 10 mM potassium phosphate in H_2O pH 7.0. The labelling reactions were quenched by the addition of an equal volume of pre-cooled 100 mM potassium phosphate pH 2.5. The labelling reactions for all time points were completed in triplicate. The quenched reaction was injected (95 µL) into a nanoACQUITY UPLC system with HDX technology (Waters Corporation). Protein digestion was performed utilising an online Enzymate immobilised BEH pepsin column (2.1 x 30 mm, Waters) at 20 °C for 1 minute with a flow rate of 150 µL/minute. Peptide separation was achieved on a C18 column (Waters Corporation Acquity UPLC BEH C18 1.7 µm, 1.0 x 10 mm) at 40 µL/min flow over 16 minutes with the following gradient; 0 min: 5% B, 7 min: 35% B, 8 min: 85% B, 11 min: 5% B, 12 min: 95% B, 13 min 5% B, 14 min: 95% B, 15 min: 5% B (mobile phases: A, water + 0.1% formic acid; and B, acetonitrile + 0.1% formic acid). Mass spectra were acquired on a SYNAPT G2-S HDMS in the positive ion mode with a *m/z* range from 290 to 2500. The mass spectrometer was operated in ToF only mode. LeuEnk peptide was used as the Lock Spray. Non-labelled peptides were identified and analysed using ProteinLynx Global Server 3.1 software (Waters), whilst DynamX 2.0 software

was used to compare the deuterium uptake-labelling rate. HDX data were mapped onto protein models using PyMOL.

Structural prediction using homology modelling

Initial homology models were produced using the Phyre 2.0 (**P**rotein **H**omology/analog**Y** **R**ecognition **E**ngine V 2.0) (43) and SWISS-MODEL online servers (44). The alignments and quality of the models were manually inspected, and overlays of the models compared in Chimera. MolProbity (45,46) was used to assess the quality of the models by analysing all atom contacts and their geometry.

1D ^{1}H NMR for glutamine amidotransferase activity

To determine whether the N-terminal domain had any glutamine amidotransferase activity, either apo- or holo-*Sc*BPL were incubated with glutamine, and the 1D ^{1}H NMR spectrum monitored for changes (i.e. decreased glutamine signals and appearance of glutamate resonances). Reactions contained 50 µM DSS, 10% D_2O, 50 mM Tris pH 8.0, 5 mM $MgCl_2$, 1 mM DTT, 20 mM L-glutamine and 1 µM apo-*Sc*BPL. To produce holo-*Sc*BPL, 50 µM biotin and 500 µM ATP were added. The pH of the samples was adjusted to 8.0 prior to NMR analysis. Standard samples of 20 mM L-glutamine and 20 mM L-glutamate in the absence of *Sc*BPL were utilised as controls, and as reference chemical shifts (47).

NMR experiments were carried out on an Agilent Inova 600 MHz spectrometer equipped with a cryo-probe at 298 K. ^{1}H signals were referenced to DSS at 0 ppm. Spectra were processed and analysed using VnmrJ 4.2 (Agilent Technologies). Spectra were collected at the following time points; 1-2 hours, 3 days, 1 week, 2 weeks and 3 weeks, with samples incubated at room temperature to allow reaction catalysis.

Results

Native MS to measure the production of purified apo- and holo-*Sc*BPL

Native nano-electrospray ionisation-mass spectrometry (nESI-MS) was first used to confirm that apo-*Sc*BPL, necessary for this study, could be purified in its non-liganded (i.e. apo) form. To produce *Sc*BPL devoid of either biotin or biotinyl-5′-AMP, cell lysate containing over-expressed *Sc*BPL was incubated with streptavidin agarose to remove excess biotin, then with saturating amounts of MgATP and a biotin accepting protein substrate, namely the *S. cerevisiae* pyruvate carboxylase biotin domain. This method has been utilised previously for the production of apo BPLs from other species (37,48) and was successfully used here to purify apo-*Sc*BPL (Figure 1 and Supplementary Figure 1). Native MS measured a deconvoluted mass of 77261 Da (calculated mass 77244 Da) (Figure 1 and Supplementary Table 1) as expected for monomeric *Sc*BPL devoid of any ligands. This was consistent with streptavidin blot analysis that also revealed the majority of *Sc*BPL had no biotinyl-transferase activity (Supplementary Figure 1). The small degree of biotinylation observed in the highly sensitive blot was not enough to influence further analytical techniques of MS and circular dichroism (CD), as it was unable to be detected by MS.

*Sc*BPL in complex with biotinyl-5′-AMP (i.e. holo-*Sc*BPL) was then produced by incubating apo-*Sc*BPL with biotin and MgATP. The reaction was then buffer exchanged into 100 mM ammonium acetate pH 6.9 to remove excess ligand and exchange the protein into the required volatile buffer for MS analysis. Ligand treatment caused the *Sc*BPL ion peaks to shift to higher *m/z* values (Figure 1) with the mass measured as 77817 Da, consistent with binding of biotinyl-5′-AMP to monomeric *Sc*BPL (expected mass 77817 Da) (Supplementary Table 1). There was no evidence of protein homodimerisation that has been observed with class I and II BPLs. A small fraction of apo-*Sc*BPL was observed, revealed by a left-side shoulder on the peaks in the holo-*Sc*BPL spectrum (measured mass 77275 Da). This apo- species is most likely the result of either ligand dissociation during buffer exchange or the stripping of ligand during ionisation despite the use of gentle MS conditions designed to maintain non-covalent interactions. MS cone voltages were increased to 200 V to improve the mass resolution of these spectra by reducing non-specific adducts for accurate MW determination. The retention of holo-*Sc*BPL at these higher voltages demonstrates the reaction intermediate interacts strongly with *Sc*BPL (Figure 1).

Apo- and holo-*Sc*BPL have similar gross structures

CD and native ion mobility-MS (IM-MS) were next employed to analyse the differences in the secondary structure and overall globular structure of *Sc*BPL when ligands bind. The CD spectra for apo- and holo-*Sc*BPL were essentially superimposable, suggesting no obvious structural differences between the two states (Supplementary Figure 2). Both the apo- and holo-*Sc*BPL contained maxima near 193 nm and minima at 208 nm and 222 nm indicative of an α-helical secondary structure, as well as an overlapping minima signal around 218 nm that indicates the presence of some β-sheet structure.

IM-MS analysis was also performed upon apo- and holo-*Sc*BPL. IM-MS measures the time an ion takes to migrate through a buffer gas, termed the drift time, which is influenced by the size and charge of the ion. Drift time can be converted into a rotationally averaged collision cross section (CCS) value that informs on the size and shape of a molecule (49). MS spectra of apo and holo-*Sc*BPL were similar, with charge state distributions between +15 and +18, and the +16 and +17 charged states equally most intense for both species (Figure 1). The CCS of apo- and holo-*Sc*BPL across all charge states were essentially the same (CCS apo-*Sc*BPL: 50.72 nm^2, CCS holo-*Sc*BPL: 50.82 nm^2) (Table 1) being within the standard error of 5% for the technique (38). These CCS values indicate there is no difference in the overall average shape or structure between apo- and ligand-bound *Sc*BPL. The CCS distribution curves (Figure 2) for apo- and holo-*Sc*BPL were similar for all charge states, reflecting normal distributions with a single apex and typically ranging from 47 nm^2 to 55 nm^2. However, the distributions for holo-*Sc*BPL were slightly narrower compared to apo-*Sc*BPL, and consistently had smaller widths at half height and standard deviations across all charge states (Supplementary Table 2). This suggests apo-*Sc*BPL may sample a greater range of spatial conformations to produce a wider variety of CCS values and, hence, may be more mobile and dynamic compared to holo-*Sc*BPL. While the CD and IM-MS analyses do not indicate large-scale structural rearrangements (i.e. overall changes in shape or domain reorganisations), these techniques are unable to detect subtle structural changes, such as loop movements or ordering that may accompany ligand binding.

Ligand binding increases *Sc*BPL stability

The stabilities of apo-*Sc*BPL and holo-*Sc*BPL were compared via both gas-phase collision induced unfolding-MS (CIU-MS) experiments and in-solution thermal denaturation assays. In the CIU-MS experiments, the MS trap energy was increased sequentially, promoting the unfolding of protein prior to ion mobility analysis. Therefore, increases in drift time resulting

from the increased conformation/size were used to monitor protein unfolding. This technique allows the identification of transition states during protein unfolding and the comparison of the stability of proteins (50). Here CIU-MS revealed ligand binding stabilises *Sc*BPL, with a higher voltage required to denature the holo-enzyme as compared to the apo protein (Figure 3. Overlays of the unfolding curves for the +17 charge state of holo-*Sc*BPL (red) and apo-*Sc*BPL (green) (Figure 3) reveal the holo-*Sc*BPL unfolding curve is right-shifted towards higher voltages by approximately 2-5 V. The magnitude of this change was comparable to other published examples of ligands binding to protein (51,52). However, the overlay of the unfolding curves for the +16 charge state of holo-*Sc*BPL (red) and apo-*Sc*BPL (green) (Figure 3) aligned. Therefore, CIU-MS suggests that ligand binding to *Sc*BPL may result in some stabilisation of *Sc*BPL, at least for the +17 charge state. Ligand binding did not appear to alter the unfolding pathway, as both apo- and holo-*Sc*BPL pass through the two short-lived transition states before being completely unfolded at similar voltages of 50-60 V (Figure 3). Together these data again demonstrate a high degree of structural homology between the apo- and holo- states. Solution-state thermal denaturation also revealed holo-*Sc*BPL is more stable than apo-*Sc*BPL, having a higher melting temperature (T_M) of 51.1 °C ± 0.04 °C compared to apo-*Sc*BPL with a T_M of 46.9 °C ± 0.3 °C ($p = 0.006$, Figure 3).

HDX-MS reveals localised ligand-induced conformational changes

HDX-MS was subsequently utilised to further investigate the structure of *Sc*BPL in both the apo- and holo-states. Here, backbone hydrogen atoms that are surface exposed, located in dynamic protein regions or not involved in the hydrogen-bonding network of the protein structure can undergo exchange to the heavier deuterium isotope when incubated in a deuterated solution. Following quenching of the exchange reaction and proteolytic digestion, the amount and sequence location of the deuterium incorporation can be detected via LC-MS by comparing the peptide masses to those measured without deuterium labelling (53-55). This provides structural information regarding the solvent accessible sites of a protein. However, these sites can only be resolved to a particular fragment peptide, and not to specific residues within it, unless peptide fragmentation is further carried out in the MS under conditions that avoid H/D scrambling. Following LC-MS of the pepsin digested HDX reactions, 149 unique peptides were detected in all replicates of the control (unlabelled) reactions of both apo- and holo-*Sc*BPL, providing 85.5% sequence coverage (Supplementary Figure 3). Deuterium incorporation within apo-*Sc*BPL was limited to defined regions, with the remaining protein having little to no deuterium uptake. This suggested the protein is primarily folded with a few surface-exposed or dynamic regions (Supplementary Figure 4). The N-terminal

extension (residues 1-220) contained multiple stretches of sequence (>15 residues in length) that did not incorporate deuterium, implying these are buried within a folded structure, supporting the presence of a structured domain within this extension. The 20 residues previously identified in the literature as being susceptible to proteases (240-260) were not present in the LC-MS coverage and therefore, confirmation of the surface accessibility of this region could not be determined from these HDX experiments. A heat map showing the difference in deuterium incorporation across the *Sc*BPL sequence between the apo and holo states was generated (Supplementary Figure 5). This highlighted that apo-*Sc*BPL incorporated more deuterium across the majority of the sequence than the holo- state (Figure 4). This suggested that the structural changes that occurred upon ligand binding reduced the ability of holo-*Sc*BPL to incorporate deuterium, consistent with the proposal that apo-*Sc*BPL has a more open or dynamic conformation and that the enzyme rigidifies upon ligand binding, consistent with the IM-MS/CCS distribution data.

Homology modelling of the class III BPL from *Saccharomyces cerevisiae*

To obtain further structural information regarding class III BPLs, homology models of the *S. cerevisiae* BPL were generated. The full sequence of the *Sc*BPL was subjected to homology modelling using Phyre 2.0 (Protein HomologY Recognition Engine) and SWISS-MODEL online servers (43,44). The results from each program were similar, with each identifying two main structured domains within *Sc*BPL. These programs also selected related proteins for the homology modelling of each structured domain, resulting in the generation of similar *Sc*BPL models. Overall, despite providing the full ScBPL sequence, only 75% of the sequence was modelled.

The Phyre program recognised that the *Sc*BPL C-terminal domain (residues 381-690) shared homology with the catalytic domain and C-terminal cap of the *Escherichia coli* BPL (22% identity), whilst SWISS-MODEL identified this region (residues 373-678) as homologous to the class I *Pyrococcus horikoshii* BPL (21% identity, 31% similarity). This was expected since all known BPLs share conserved structural folds of the catalytic domain and the C-terminal cap necessary for protein biotinylation (28,56). Phyre modelled the *Sc*BPL catalytic domain using the *E. coli* BPL structure PDB: 2ewn (11), whilst SWISS-MODEL utilised the *P. horikoshii* BPL structure PDB: 2e10 (Figure 5A and B respectively). The model produced by SWISS-MODEL was dimeric, likely because of the dimeric nature of the *Ph*BPL template. However, the native MS described above demonstrated *Sc*BPL is monomeric (Figure 1). Therefore, only one of the protein chains was utilised when assessing this model. The

secondary structural elements of the two models overlaid well (RMSD of 1.411 Å) (Figure 5C), however, there was significant variation in the connecting loop regions that lacked secondary structure (RMSD: 8.506 Å between the full models). Whilst SWISS-MODEL attempted to model these residues, Phyre often truncated the loops, not modelling the full sequences in these regions. MolProbity (45,46) was utilised to assess the quality of the structures by evaluating whether residues adhered to appropriate ramachandran angles (Figure 5A and B). 96% of residues in the Phyre model were in allowed regions of Ramachandran space (89% classified as favoured), with 10 outlier residues (Supplementary Figure 6). In comparison, the SWISS-MODEL model placed 95% of residues in allowed regions (83% in favoured regions), with 15 outliers (Supplementary Figure 7). The outliers for both models were mostly located in loop regions that are less likely to be modelled correctly due to difficulty in *de novo* modelling for loops not present in the homologous template protein. These models aligned well with the catalytic domains from crystal structures of class I/II BPLs, with small RMSD values (Supplementary Tables 3 and 4). Furthermore, sequence motifs and catalytic residues that are important for BPL function, including the GRGRXG and KWPND motifs along with a key tryptophan residue (W430 in *Sc*BPL) required for π-π stacking interactions with ATP, were conserved and correctly positioned within the models of the catalytic domain, when compared to known BPL structures.

Each program also modelled a structured domain between residues 1 to 220 of the N-terminal extension of *Sc*BPL. The location of this domain is compatible with previous limited proteolysis experiments that predict a folded domain in the first 240 residues of *Sc*BPL (19). Both Phyre and SWISS-MODEL predicted the N-terminal 220 residues form a domain that shared homology with the glutamine amidotransferase (GATase) fold. Phyre determined the most homologous structure for this N-terminal extension was the GATase subunit (PDB: 2ISS) (57) from the *Thermatoga maritima* pyridoxal 5′-phosphate synthase complex (Figure 6A), with 23% sequence identity between these two proteins (c.f. 22% identity between the *Sc*BPL catalytic domain and the *E. coli* BPL catalytic domain it was modelled on). The other top homology results identified by Phyre for this *Sc*BPL region were other GATases from other enzyme complexes and/or species. Similarly, SWISS-MODEL utilised a GATase to model the N-terminal domain, this time using the GATase from the imidazole glycerol phosphate synthase complex from *T. maritima* (hisH, PDB: 1KXJ) (17% identity, 28% similarity) (58) (Figure 6B). Both models generated consistent tertiary structural folds with core secondary structure elements aligning well (RMSD 1.761Å) (Figure 6C). As described above, structure predictions of disordered or loop regions were more varied between the two

models (RMSD 4.480 Å across the length of the whole models), likely due to difficulties in correctly modelling these features. Once again, Phyre truncated the predicted loops, whilst SWISS-MODEL attempted to model all loops. The Phyre model contained 95% of residues in the allowed Ramachandran regions (95% in the favoured region), with only 8 outliers (Figure 6A and Supplementary Figure 8). The SWISS-MODEL model had similar results with 97% of residues in the allowed Ramachandran regions and 86% in the favoured region but with 6 outliers (Figure 6B and Supplementary Figure 9). Just as for the C-terminal domain, most of the outliers were located in loops or less structured regions that are less likely to be correctly modelled. Unexpectedly, a catalytic triad of residues responsible for the enzymatic function of GATases (cysteine, histidine and glutamic acid) was found to be conserved and correctly positioned in both models of the N-terminal domain of *Sc*BPL (residues C89, H215, E217) (Supplementary Figure 10A and B). These residues are involved in catalysing the conversion of glutamine to glutamate and ammonia in GATases, but their relevance here in BPL is unknown.

Neither program found structural homologues for the 160 residues located between residues 221-380 of *Sc*BPL. Consequently, these residues remained unmodelled. This region of *Sc*BPL was located around the previously identified protease susceptible site (residues 240 to 260) (19). Thus, homology models of the N and C-terminal domains of *Sc*BPL have been produced, but how the intervening residues connect them is unknown. Nevertheless, the models can be utilised for understanding the structural implications of the previously recorded HDX data. Ultimately the models produced by Phyre were utilised as they did not attempt to model the variable loop regions, regions that are more likely to be disordered and/or mobile.

_Sc_BPL N-terminal domain does not function as a glutamine amidotransferase (GATase)

Homology modelling predicted a glutamine amidotransferase structural fold for the N-terminal domain. We therefore set out to ascertain whether *Sc*BPL has any corresponding GATase activity. GATase function is dependent on a catalytic triad of residues (cysteine, histidine and glutamic acid) that are responsible for catalysing the conversion of glutamine to glutamate and ammonia. The ammonia produced is normally then transported through the full GATase protein complex to other enzymatic subunits for further catalytic reactions (59,60). The catalytic triad residues are conserved in the N-terminal domain of *Sc*BPL (residues C89, H215, E217) (Supplementary Figure 10A and B), and the homology models predicted them to be appropriately positioned for catalytic activity to be possible. However, addition of glutamine

(and 5 mM $MgCl_2$ to aid GATase catalysis) to *Sc*BPL did not result in production of glutamate, even after two weeks at room temperature, as determined by ^{1}H 1D NMR (Supplementary Figure 10C). In addition, the biotinylation capability of *Sc*BPL was unaltered in the presence of glutamine (data not shown), demonstrating that the conventional GATase domain substrate (glutamine) had no influence on the activity of *Sc*BPL. Therefore, we concluded that *Sc*BPL does not contain GATase activity, despite having a GATase homology domain.

HDX experiments support the homology model structures of *Sc*BPL N-terminal and catalytic domains

The HDX rates for apo-*Sc*BPL were mapped onto the PHYRE models to determine if predicted surface exposed regions were corroborated by the experimental data (Figure 7). The results for apo-*Sc*BPL were used to assess the structure of *Sc*BPL prior to ligand binding. This species had the greatest deuterium incorporation rates, suggesting it was the most open and dynamic conformation of *Sc*BPL. It should be noted that the resolution of the technique allows the location of deuterium incorporation to be localised to a short peptide (up to 25 residues in length), rather than individual residues. Regions of the catalytic domain with greater deuterium incorporation were located on predicted exposed surfaces, loop regions and in the active site (Figure 7B and Supplementary Figure 4), and there was no deuterium incorporation observed in the inner protein core regions. Specifically, deuterium incorporation occurred in the active site ligand binding pockets, including within β-sheets β2, β3, β5 and β6 that form the back surface of the active site pockets along with the biotin-binding loop (residues 416-437), part of the adenylate-binding loop (residues 552-565) and another disordered loop that helps form the biotin-binding pocket (residues 387-393). These data are consistent with other BPLs where analogous loops are disordered in crystal structures of the apo-enzyme, but become ordered upon ligand binding. Other predicted surface loops from the models of *Sc*BPL (residues 495-528 and part of β5 that lies within a core β-sheet, residues 604-611 and the adjacent ends of α4 & α5 that this linker joins, and residues 443-460 together with part of α2 that the loop connects to) also incorporated deuterium, compatible with the flexible structure proposed by the model. These latter two linkers are the only sites of deuterium uptake found on the face opposite the ligand-binding site. There is also evidence of deuterium incorporation across the entire C-terminal cap domain (comprised of surface exposed β-strands) and the flexible termini of the protein.

Few areas in the C-terminal catalytic domain were not included in the LC-MS coverage and, therefore, the HDX uptake in these regions is unknown. This included sections of β3 and β6

that form the part of the β-sheet protein core, and part of α2, α5 and α6 that form the face of the domain opposite to the ligand-binding site. There were also areas that appeared surface exposed within the model that did not incorporate deuterium, such as the remainder of the adenylate-binding loop (residues 566-578) and the following surface residues linking it to α4 of the domain's core (residues 578-587), along with the loop linking the core β5 and β6 strands (residues 538-545). The absence of deuterium exchange in these seemingly surface accessible regions, particularly the parts of the domain face opposite to the ligand-binding site that were included in the LC-MS coverage, suggested these regions have reduced solvent accessibility or are involved in hydrogen bonding networks such that the backbone hydrogen atoms are unable to undergo exchange. This may be due to interactions with the N-terminal domain or 160-residue linker region, which may bury these surfaces, preventing deuterium incorporation.

Deuterium uptake was also mapped onto the Phyre model of the N-terminal domain (Figure 7A and Supplementary Figure 4). Interestingly, whilst many regions of this small domain incorporated deuterium, they only underwent very low levels of exchange, consistent with a folded domain structure. These regions of low deuterium incorporation included the secondary structure elements located on the surface of the model, including two pairs of surface-located anti-parallel β-sheets (β5 & β8, and β6 & β7) along with α1, β2 and α4. Predicted disordered loops, some of which were not modelled (Figure 7) also incorporated deuterium, including the loops between β1-α1 (residues 6-16), α1-β2 (residues 26-30), α2-β3 (residues 47-52), β3-α3 (residues 55-60), β5-β6 (residues 127-138, incorporates a small helical element), β6-β7 (residues 143-151), β7-β8 (residues 156-161), β8-β9 (residues 165-171) and β9-β10 (residues 176-198). The regions of strongest deuterium exchange were α2, α3, α5, α6 and β6, all of which are surface located.

Few regions in the N-terminal domain were also not covered by LC-MS analysis, hence, the HDX uptake in these regions could not be determined. This included the N-terminal half of α4, the loop between α5 and β5 (residues 95-121) and part of the β-sheet core (parts of β4, β9 and β10). There were also two regions of the N-terminal domain that were excluded from deuterium uptake including most of the core β-sheet that was covered by LC-MS and the C-terminal half of the surface exposed α4 that was also included in the LC-MS coverage. The absence of deuterium incorporation in these seemingly exposed regions may again be due to these surfaces being utilised for interactions with the remainder of the BPL enzyme, thereby occluding solvent. Together there was excellent agreement between the HDX data and the

predicted model of the N-terminal domain. In addition, the data suggested a potential interaction interface between the N-terminal domain and the catalytic domain.

Incorporation of deuterium in the unmodelled 160-residue linker between the N-terminal and catalytic domains was analysed. The majority of this region incorporated either small amounts of deuterium after longer incubation times (greater than 10 minutes), or no deuterium at all (Figure 7C and Supplementary Figure 4). Overall there was less HDX in this region than in either the catalytic domain or N-terminal extension, with fewer patches of lower levels of deuterium uptake. Together these results suggest this region is relatively immobile and/or buried.

Overall, the HDX validated the proposed structure for *Sc*BPL, revealing that the protein is mostly folded with a structured domain in the N-terminal extension that is linked to the catalytic domain by a folded linker. The N-terminal and catalytic domains are suggested to interact with each other through the opposite face of the catalytic domain to the ligand binding site, the loop between β5 - β6 and/or the remainder of the adenlyate binding loop in the catalytic domain, and the α4 helix in the N-terminal domain. We propose these seemingly surface exposed regions are occluded from deuterium exchange, possibly due to protein interactions with the remainder of the BPL that prevent solvent accessibility.

HDX and homology modelling reveal locations of ligand-induced conformational changes

Just as the raw levels of deuterium incorporation for apo-*Sc*BPL were mapped onto the homology models, the difference in the deuterium incorporation between apo- and holo-*Sc*BPL was mapped onto these structures to identify regions of *Sc*BPL that undergo structural alteration upon ligand binding (Figure 8). Overall, there was a general reduction in the incorporation of deuterium by *Sc*BPL following ligand addition (Figure 4). This is consistent with the enzyme becoming more rigid, thereby reducing surface exposure and preventing flexible regions from undergoing deuterium exchange. As expected, regions surrounding the ligand binding sites had reduced deuterium uptake following ligand binding (Figure 8B and Supplementary Figure 5). These regions included the loop that helps form the biotin-binding pocket located between β1 and α1 (residues 387-397), along with part of the biotin-binding loop (residues 431-437 from the loop 416-437) (Supplementary Figure 11). Residues 549-565, from the partially modelled adenylate-binding loop, as well as parts of the β-sheet that

form the base of the ligand-binding pocket (particularly β2, β3, β5 and β6), also had reduced deuterium uptake in holo-*Sc*BPL.

Regions of the N-terminal domain with reduced deuterium uptake due to ligand binding were located on a single surface patch on one side of the protein model (Figure 8A and Supplementary Figure 5). This was the same side of the domain that had the greatest levels of deuterium exchange in the apo-*Sc*BPL structure. The regions of the largest difference in deuterium incorporation between apo- and holo-*Sc*BPL were residues 54-68, located on β3, α3 and the interconnecting loop, and residues 90-97 from α5 (Supplementary Figure 11). Finally, the linker region (residues 220-380), for which no homologous structure was identified for modelling, generally showed no greater deuterium incorporation in either state (Figure 8C and Supplementary Figure 5). These HDX data together with homology modelling of *Sc*BPL provided details on areas of localised conformational changes that occur with ligand binding.

Discussion:

Class III BPLs, including that from *S. cerevisiae,* contain a unique N-terminal extension for which there is little information available regarding its function, and even less about its structure. The absence of an atomic resolution structure of a class III BPL has hindered the understanding of the purpose of this unique region and its role in biotinylation. Here we utilised molecular modelling and structural mass spectrometry techniques to build and validate a model for the class III enzyme *Sc*BPL. We also investigated conformational changes that occur with the binding of the ligands biotin and MgATP. Whilst IM-MS and CD spectroscopy did not identify any large-scale structural rearrangements, HDX and IM-MS did reveal holo-*Sc*BPL had reduced flexibility and/or mobility consistent with stabilisation of ligand binding loops and core structures. This is consistent with the increased stability observed for holo-*Sc*BPL that was identified by thermal denaturation assays and CIU-MS of the +17 charge state of *Sc*BPL. HDX-MS allowed these smaller ligand-induced structural rearrangements to be mapped upon our model, revealing they were situated around the active site on the catalytic domain where the ligands bind. Changes were also located on a surface patch on the homology model of the N-terminal domain (discussed further below). These conformational changes are localised structural rearrangements such as loop or helix rigidification or a reduction in protein dynamics, rather than a reorganisation of the whole enzyme, explaining why the lower-resolution techniques that assess the conformationally averaged overall structure (IM-MS & CD) did not identify them.

Homology models of the N-terminal and catalytic domains of *Sc*BPL were produced to allow a structural understanding of *Sc*BPL, and to aid analysis of the HDX results. Utilising the full-length *Sc*BPL sequence produced individual structures of both the N-terminal and catalytic domains of *Sc*BPL. The *Sc*BPL C-terminal region was homologous to, and subsequently modelled based on, the catalytic domain and C-terminal cap of other species BPLs. This is not surprising, considering the catalytic domains between different BPLs are well conserved to maintain the biotinylation function (28,56). The N-terminal domain was predicted to have homology with and, therefore, a fold analogous to the structure of a glutamine amidotransferase (GATase). Overlaying the deuterium incorporation data with these models provided verification that the predicted surface exposed regions corresponded to the experimental data. These analyses of the HDX data were consistent with the predicted folds for each *Sc*BPL domain. Whilst the N-terminal 240 and C-terminal 380 residues of *Sc*BPL were modelled in individual domains, the interconnecting 160-residues (220-380) had no homologous structure identified and, thus, were not modelled. The relatively low HDX in this

region suggested this connecting linker is folded and/or partially buried. We hypothesise this region forms additional structured elements (α-helices and β-strands) that link the two domains together. The HDX data together with the homology models also revealed seemingly surface exposed regions that were not susceptible to deuterium uptake. This involved a surface helix in the N-terminal domain, and part of the backside of the catalytic domain (opposite side to the ligand binding side) along with the C-terminal part of the adenylate-binding loop and its linker back to the protein core, and the loop joining β5 and β6. These sites are hypothesised to form interaction surfaces with the remainder of the BPL, such that the surface is covered by protein and occluded from solvent to prevent deuterium uptake. The N- and C-terminal domains have not been successfully produced as individual proteins, preventing their use as controls for mapping the interaction interface between the two domains. HDX of the individual domains would reveal surface exposed regions that then become buried when the remainder of the enzyme is present, revealing these surfaces are involved in intra-molecular interactions. Hence, further efforts to produce soluble, folded proteins of the individual domains would be beneficial. The determination of an atomic resolution structure of a class III BPL, particularly that from *S. cerevisiae,* is required to enable confirmation of the structural hypotheses proposed above (folded, buried linker and interaction surfaces of each domain). HDX and IM-MS data generated here can also serve as constraints for further modelling of the N- and C-terminal domains, and docking of how the two domains interact.

The identified homology between the N-terminal domain and a GATase was further investigated, as GATase-like structures and function have not previously been linked to biotinylation. GATases are normally situated in multi-subunit protein complexes where they convert glutamine to glutamate and ammonia, before funnelling the ammonia to other subunits in the complex for metabolite synthesis (59,60). Activity assays revealed *Sc*BPL does not contain this catalytic activity despite the models predicting the *Sc*BPL N-terminal domain maintains homologous correctly position catalytic residues as a GATase. There is no obvious reason for GATase function in eukaryotic BPLs, with no requirement for ammonia or glutamate for biotinylation. Hence, it is not clear why the enzyme may contain a similar fold to a GATase. It remains to be investigated whether the N-terminal domain has any affinity for binding glutamine, the substrate of GATases. Mutation of the GATase catalytic residues that are conserved in *Sc*BPL could be utilised in activity assays to delineate whether these residues are important for biotinylation or just conserved to maintain the structure of the GATase-like fold predicted for this domain.

HDX was also utilised in this study to investigate the subtle structural changes that occur within *Sc*BPL and result in an increase in stability and a reduction in proteolysis (previously reported (19)). The differences in deuterium incorporation between apo- and holo-*Sc*BPL were mapped onto the homology models to understand the structural context of the HDX-identified conformational changes. The ligand-induced conformational changes localised in the active site of the catalytic domain can be understood using our knowledge of other species BPLs. Firstly, the binding of biotin and MgATP fill the active site to prevent solvent entering this pocket and also occlude the active site residues from solvent or involve them in bonding interactions with the reaction intermediate. This explains the reduction in HDX in the active site pocket when ligands bind. Secondly, the surface loops in the active site that have reduced HDX following ligand binding are analogous to the biotin-binding loop (residues 418 to 434 in *Sc*BPL, residues 116-124 in *Ec*BPL) and the adenylate-binding loop (residues 555 to 578 in *Sc*BPL, residues 212-234 in *Ec*BPL) of other BPLs. Structural studies of other species BPLs revealed these loops are disordered in the apo-enzymes and that biotin binding induces the ordering of the biotin-binding loop to correctly position residues of the ATP binding pocket that allows ATP to bind (11,13). Following ATP binding and reaction intermediate formation, the adenylate-binding loop then also becomes ordered and folds over the active site to protect the reaction intermediate from dissociation and hydrolysis by solvent (11,14). The ordered-binding of biotin then ATP, and disordered-to-ordered transitions of these active site loops, are predicted to be maintained in *Sc*BPL, with the residues in these loops having reduced deuterium uptake following ligand binding – implying structural rearrangement to reduce disorder/dynamics and become ordered, and involvement in interactions with the ligand. This stabilisation of loops and recruiting loop residues for interactions with ligands can also explain the reduced deuterium uptake in the third loop identified, which forms part of the active site by forming part of the biotin binding pocket (residues 386-395).

Small structural rearrangements induced by ligand binding were also identified in the N-terminal domain. These allosteric changes formed a small surface patch that became more structured, ordered or buried upon ligand binding such that deuterium incorporation was reduced. The function of this surface is unknown but it is predicted that it could become rearranged to provide a surface to interact with either the remainder of the BPL, the biotin domain substrate, or both. Previous studies provide evidence of interactions between the N-terminal extension and the biotin domain that help to form an enduring complex for protein

biotinylation to proceed (12,21,22). These ligand-induced structural changes could also occur to expose residues to contribute towards assisting the catalytic domain and helping catalyse the biotinylation reaction. Analysis of crucial residues for catalysis in other classes of BPLs, and determining whether these are conserved in the catalytic domains of class III BPLs could help determine whether the N-terminal domain is necessary to contribute residues to aid with reaction catalysis. Mutagenesis of this N-terminal domain surface, or a high-resolution structure of a BPL-biotin domain complex, would provide vital information into the function of this surface and the relevance of the ligand-induced rearrangements for biotinylation.

These HDX experiments support and help refine the structural hypothesis proposed from the previous limited proteolysis experiments that first identified conformational differences between apo- and holo-*Sc*BPL (19). In these experiments, proteolysis occurred between residues 240 and 260, separating the N-terminal extension from the remainder of the BPL that contains the catalytic domain. Proteolysis of this 20-residue 'linker' was reduced when *Sc*BPL was incubated with biotin and MgATP, leading to the hypothesis that ligand binding introduced constraining conformation changes to prevent proteolysis in this otherwise surface exposed, flexible linker (19). The present study has allowed us to refine this original model of *Sc*BPL structure. HDX and homology modelling confirmed a structured domain in the N-terminal extension located within the first 220 residues. This study also revealed a larger, folded, mostly buried linker of 160 residues separates this N-terminal domain from the catalytic domain that initiates at residue 381. The 20-residue protease susceptible region (residues 240-260) may still form a surface exposed loop within the mostly folded, buried 160-residue linker such that it is accessible to proteases. The lack of HDX coverage over this 20-residue site prevents confirmation of this hypothesis.

The *Sc*BPL models and structural information produced here can be utilised to understand the mechanism behind N-terminal mutations that cause the disease MCD. Two common N-terminal domain mutations are L216R and L237P (32,61,62), which correspond to residues A52 and S79 in the N-terminal domain of *Sc*BPL. Previous studies have shown the mutations reduce the affinity of the human BPL for the biotin domain (12), however, the mechanism by which they do this has not been determined. A52 (equivalent to L216 in the human BPL) is located at the beginning of the core-located β3 strand of the N-terminal domain, whilst S79 (equivalent to human L237) is located at the C-terminal end of α4, on the core-directed surface with its side chain directed towards the N-terminal domain interior (Figure 9). According to our homology model, mutation L216R (A59 in *Sc*BPL) would introduce a

positively charged side-chain into the core of the N-terminal domain, a change that is expected to be highly disruptive for domain folding. Replacement of S79 with proline would disrupt proper formation of helix α4. Interestingly, helix α4 is seemingly surface located, however, the C-terminal half of it that was included in the LC-MS coverage did not incorporate deuterium. It was therefore identified as a potential interaction interface with the remainder of the BPL. Hence, these mutations are likely to disrupt the secondary and tertiary structure of the N-terminal domain, and possibly the overall BPL structure, such that it can no longer interact with the biotin domain, leading to the pathogenic loss of BPL activity.

These data presented here reveal novel structural insights to *Sc*BPL by proposing structural models of the N-terminal and catalytic domains within *Sc*BPL, and identifying the conformational changes that occur within the ligand binding site and N-terminal domain upon ligand binding. We now seek to further validate these models using further empirically derived data from higher-resolution structural technique. This will help reveal how the two modelled domains are orientated relative to each other and ultimately provide a structure of the complete enzyme. Further pursuing X-ray crystallography, or alternative structural techniques of cryo-electron microscopy or the emerging micro-electron diffraction, may enable the determination of an atomic resolution structure.

Acknowledgments

The authors would like to acknowledge the Biophysical Characterisation Facility at the University of South Australia Cancer Research Institute where the circular dichroism experiments were completed.

Funding

This work was supported by the National Health and Medical Research Council of Australia (application APP1068885). LMS is the recipient of an Australian Government Research Training Scheme stipend and an ASBMB Fred Collins Fellowship.

Author Contributions

LMS and SWP conceived the project; LMS, TLP, KP collected the data; LMS, TLP, KP, PB, KLW and SWP analysed the results; LMS, TLP, KP, PB, GWB, KLW and SWP assisted with manuscript preparation.

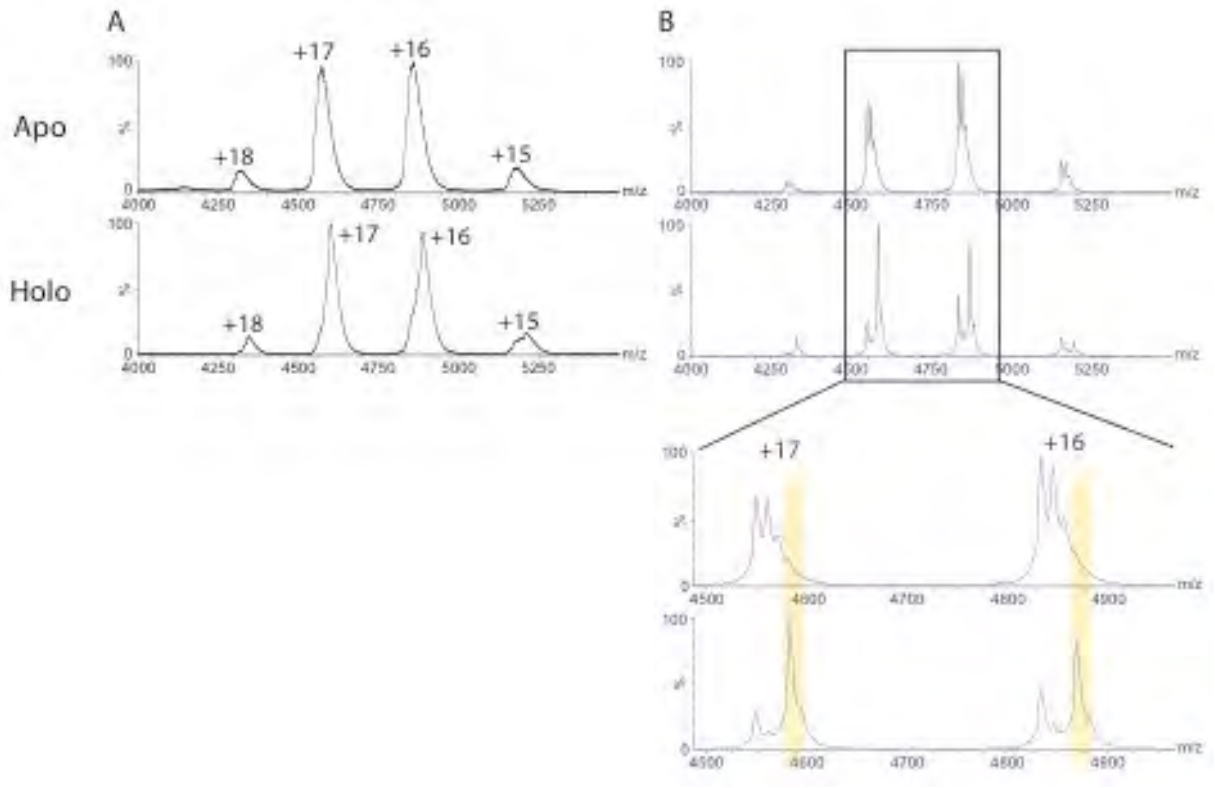


Figure 1. Native MS can distinguish between apo- and holo-treated *Sc*BPL. Spectra at cone voltage of A) 50 V and B) 200V (blue). The inset in panel B shows a zoomed in view of charge states +17 and +16, with yellow highlights indicating the mass range expected for the holo species.

Table 1. Collision cross sections (CCS) reveal no difference in the overall structures of apo- and holo-*Sc*BPL.

	CCS (nm^2)				
	15+	**16+**	**17+**	**18+**	**AVE CCS ± SEM (nm^2)**
apo	50.85	49.99	50.49	51.54	50.72
holo	50.85	50.39	50.49	51.54	50.82

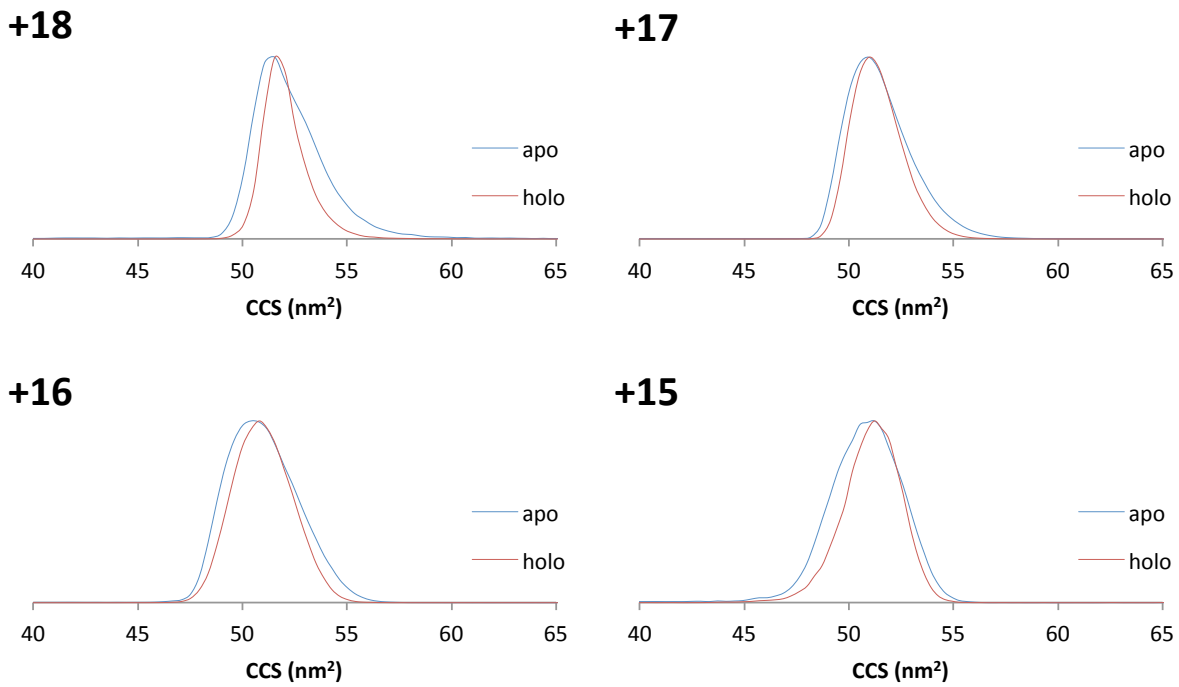


Figure 2. CCS distributions of apo- (blue) and holo-*Sc*BPL (red) across the four charge states observed by IM-MS. The standard error of the IM-MS technique is 5% (38).

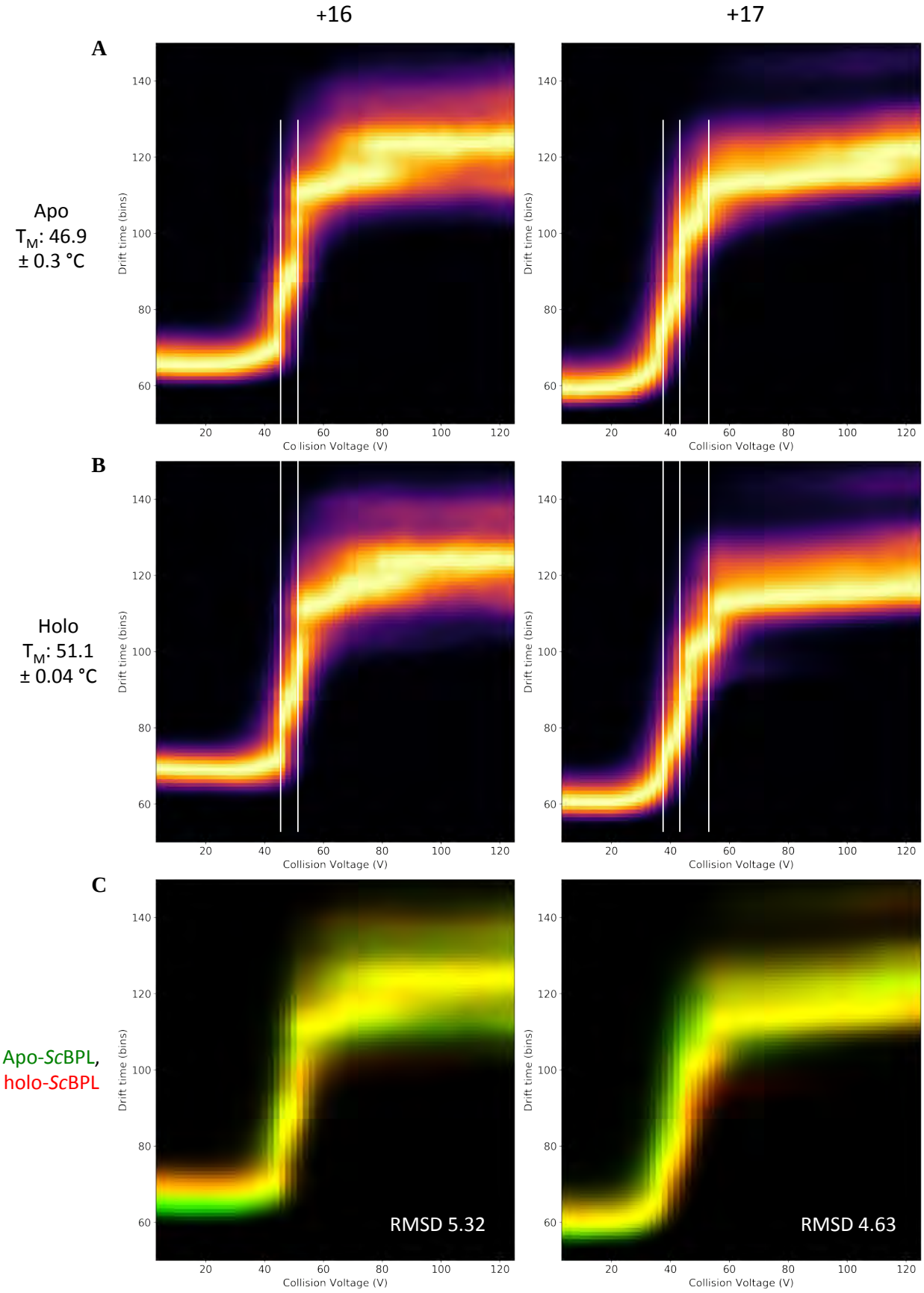


Figure 3. Holo-*Sc*BPL is more stable than apo-*Sc*BPL in solution thermal denaturation assays (T_M annotated on left of figure) and collision-induced unfolding MS (CIU-MS). CIU-MS

completed on the two most intense charge states, +17 and +16 for A) apo-*Sc*BPL and B) holo-*Sc*BPL. Increasing voltages (x-axis) are applied to the protein to unfold it, causing an increase in protein size and, therefore, drift time (y-axis). White lines mark the voltage at the end of each transition state for apo-*Sc*BPL to compare with holo-*Sc*BPL. C) Overlays of the unfolding curves for apo- (green) and holo-*Sc*BPL (red) are included.

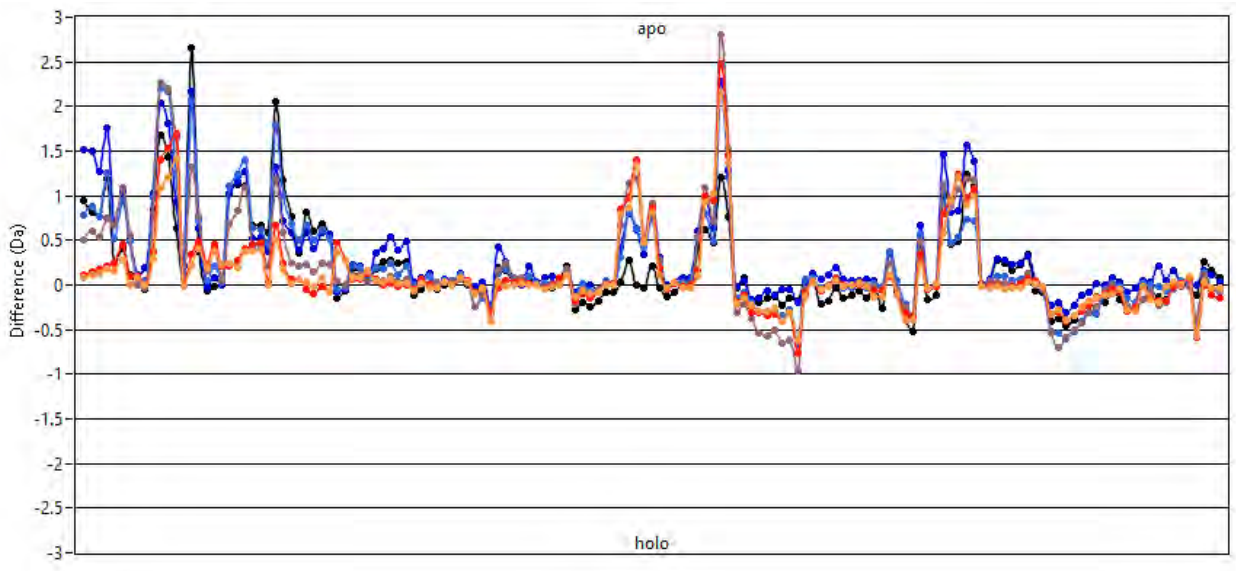


Figure 4. Apo-*Sc*BPL incorporated more deuterium across majority of the *Sc*BPL sequence than holo-*Sc*BPL. The difference in the uptake of deuterium between apo- and holo-*Sc*BPL for each peptide after 30 seconds (orange), 1 minute (red), 10 minutes (grey), 30 minutes (light blue), 1 hour (dark blue) and 4 hours (black) exposure to deuterium.

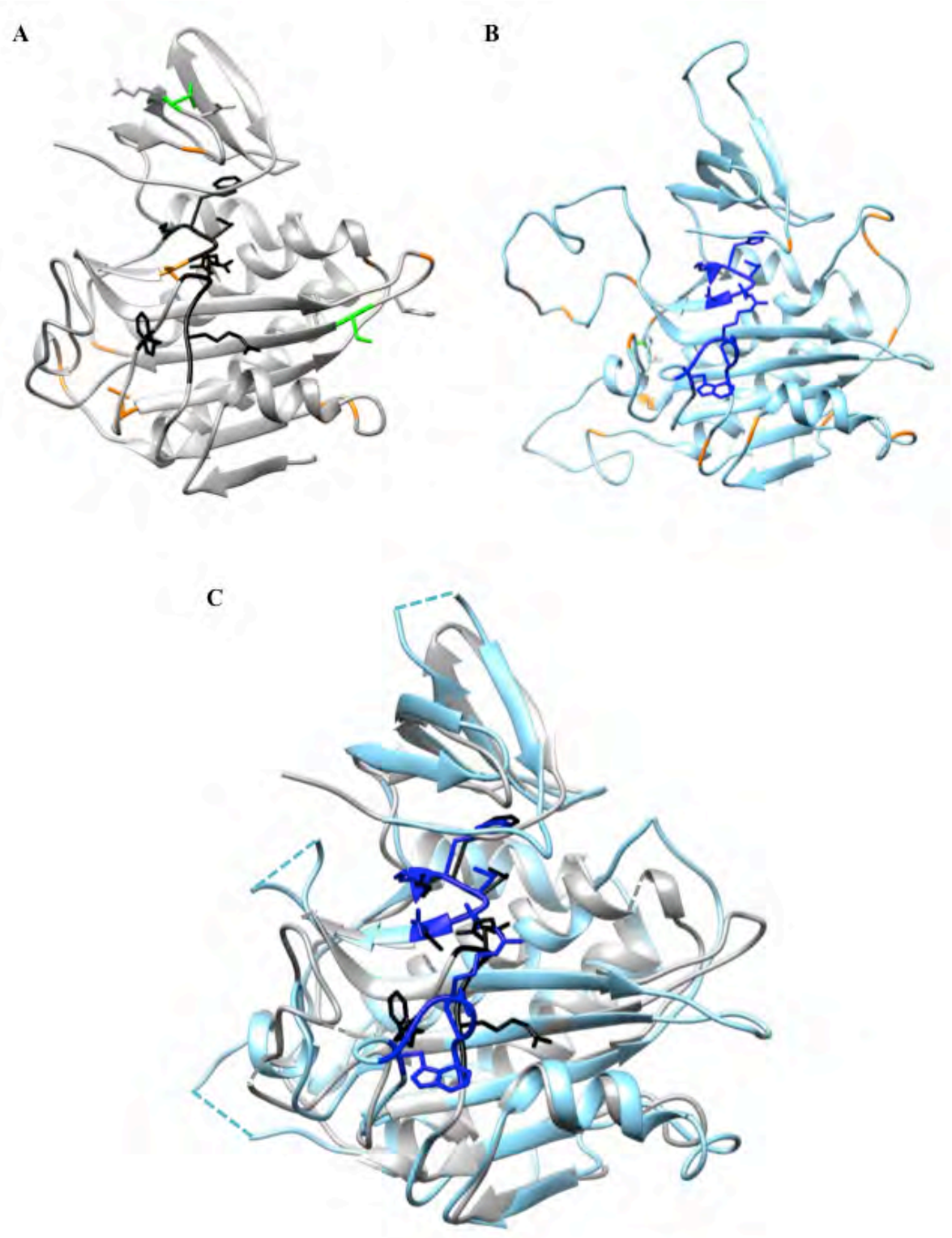


Figure 5. Homology models of the *Sc*BPL catalytic domain and C-terminal cap produced by A) Phyre (grey) and B) SWISS-MODEL (blue). Important conserved motifs and residues are highlighted in black and dark blue for the models respectively. Residues that are ramachandran outliers (orange) or have cis isomerisation (green) are also annotated. C) An overlay of the core structure of the models, with large surface loops truncated and the unmodelled regions of more than 1 residue removed, reveals the similarity in the models produced by Phyre and SWISS-MODEL.

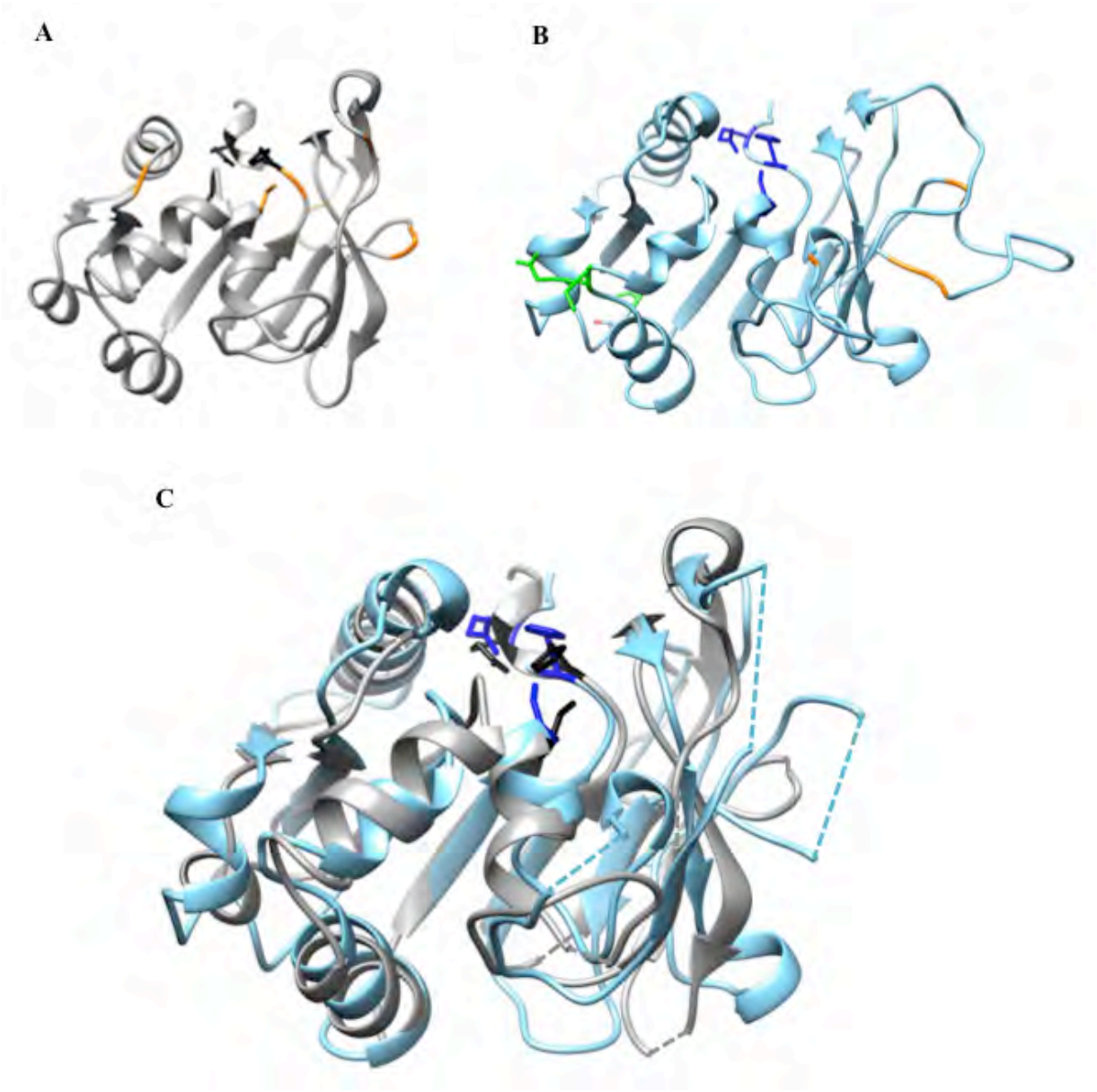


Figure 6. Homology models of the *Sc*BPL N-terminal domain generated by A) Phyre (grey) and B) SWISS-MODEL (blue). The conserved GATase catalytic residues (C89, H215, E217) are highlighted in black and dark blue for the models respectively. Residues that are ramachandran outliers (orange) or have cis isomerisation (green) are annotated. C) An overlay of the core structure of these two N-terminal models, without the large disordered surface loops and the unmodelled regions of more than 1 residue removed, reveals similarity in the models produced by both modelling programs.

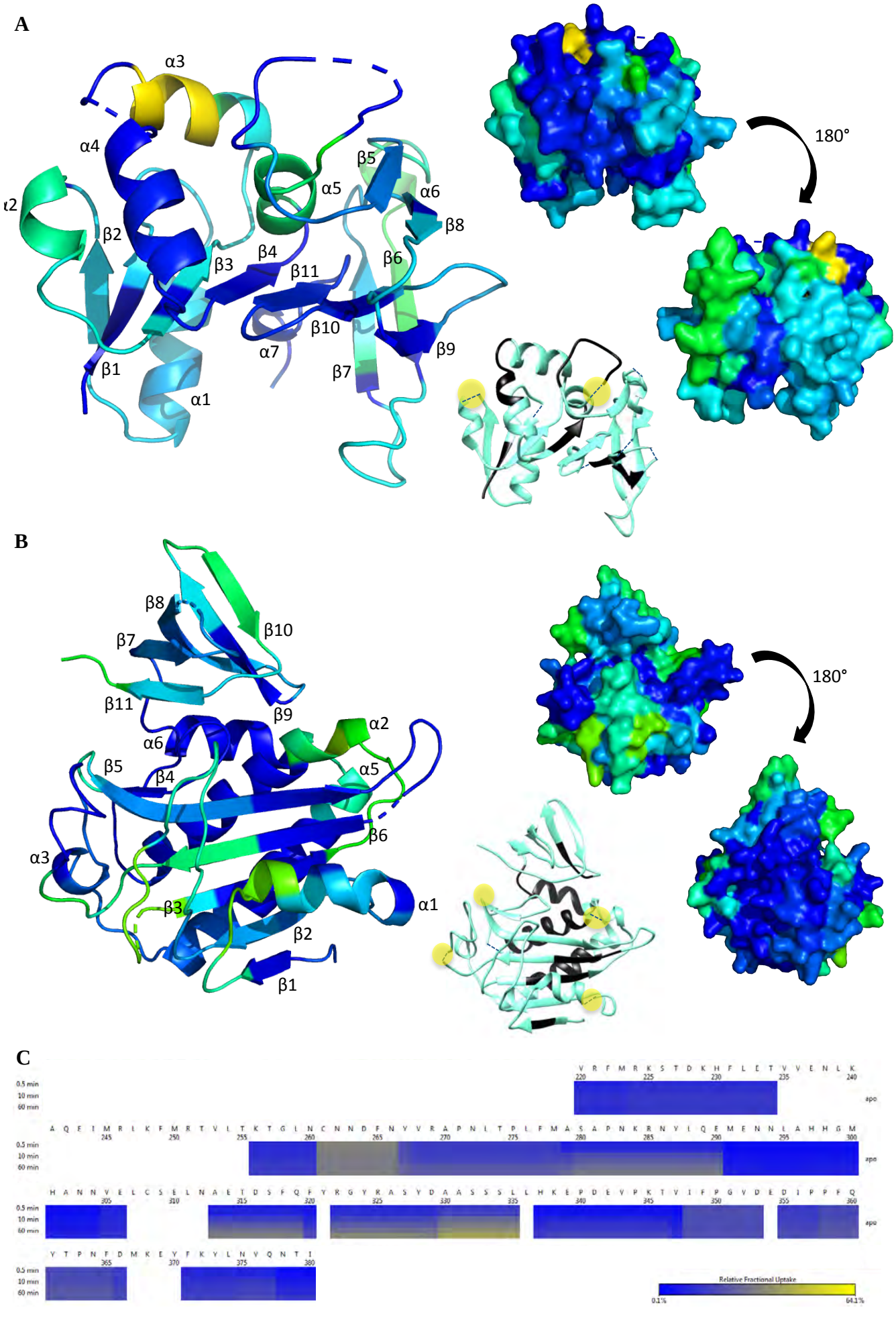

A
α3
α4
α2
β2
β3
β4
α5
β5
α6
β8
β6
β11
β10
β9
β7
α7
β1
α1
180°
B
β8
β7
β10
β11
β9
α2
α6
β5
β4
α5
β6
α3
β3
α1
β2
β1
180°
C
0.5 min
10 min
60 min
apo
Relative Fractional Uptake
0.1%
64.1%

Figure 7. Deuterium incorporation for apo-*Sc*BPL after 1 hour mapped onto the A) N-terminal domain and B) catalytic domain models, in both ribbon and surface fill representations. A rainbow colour scheme indicates the level of deuterium uptake, with dark blue showing no exchange through to light blue, green and yellow representing increasing levels of deuterium incorporation. A ribbon representation of each domain (aqua green) shows regions not included in the HDX LC-MS coverage (black) and linker regions that were not modelled but incorporated deuterium (highlighted yellow). C) A heat map of deuterium incorporation across the unmodeled linker region is included. Here, yellow colouring indicates deuterium incorporation, while blue indicates unexchanged regions.

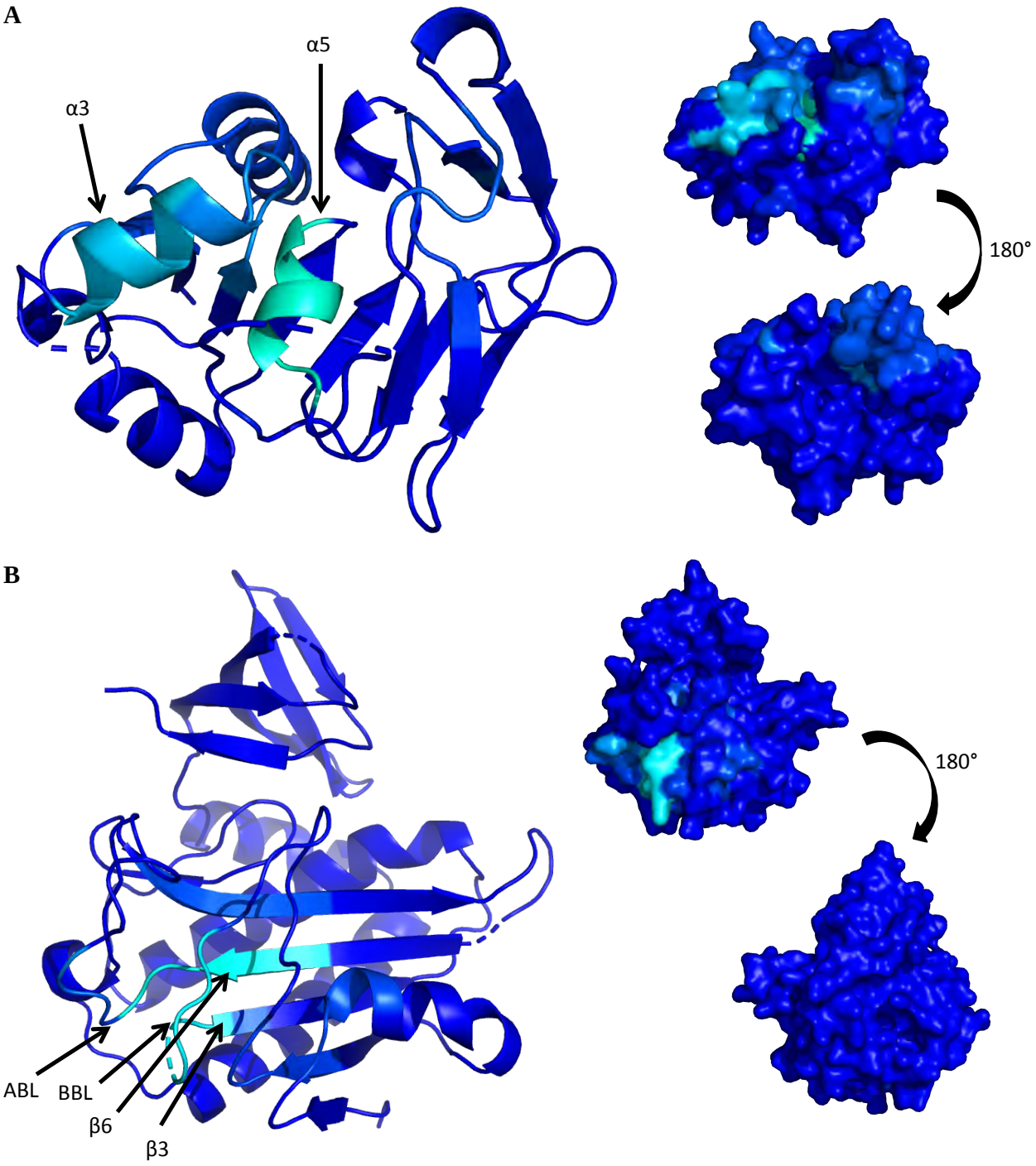

A
α3
α5
180°
B
180°
ABL
BBL
β6
β3
C
0.5 min
10 min
60 min
apo - holo
Relative Fractional Uptake apo - holo
-33.8%
33.8%

Figure 8. Sites of the A) N-terminal domain and B) C-terminal catalytic domain of *Sc*BPL that have altered deuterium uptake as a result of ligand binding. Results from the difference in deuterium uptake of apo-*Sc*BPL and holo-*Sc*BPL after 1 hour are mapped onto the N-terminal and catalytic domain models for both ribbon and surface fill representations. Dark blue colouring indicates no difference in deuterium incorporation between apo- and holo-*Sc*BPL, whilst light blue and green indicate increasing amounts of deuterium incorporation in apo-*Sc*BPL compared to holo-*Sc*BPL. C) The heat map showing the differences in deuterium exchange for the linker region (residues 220-380) is also included. Red represents greater HDX for apo-*Sc*BPL, whilst blue indicates more deuterium incorporation in holo-*Sc*BPL.

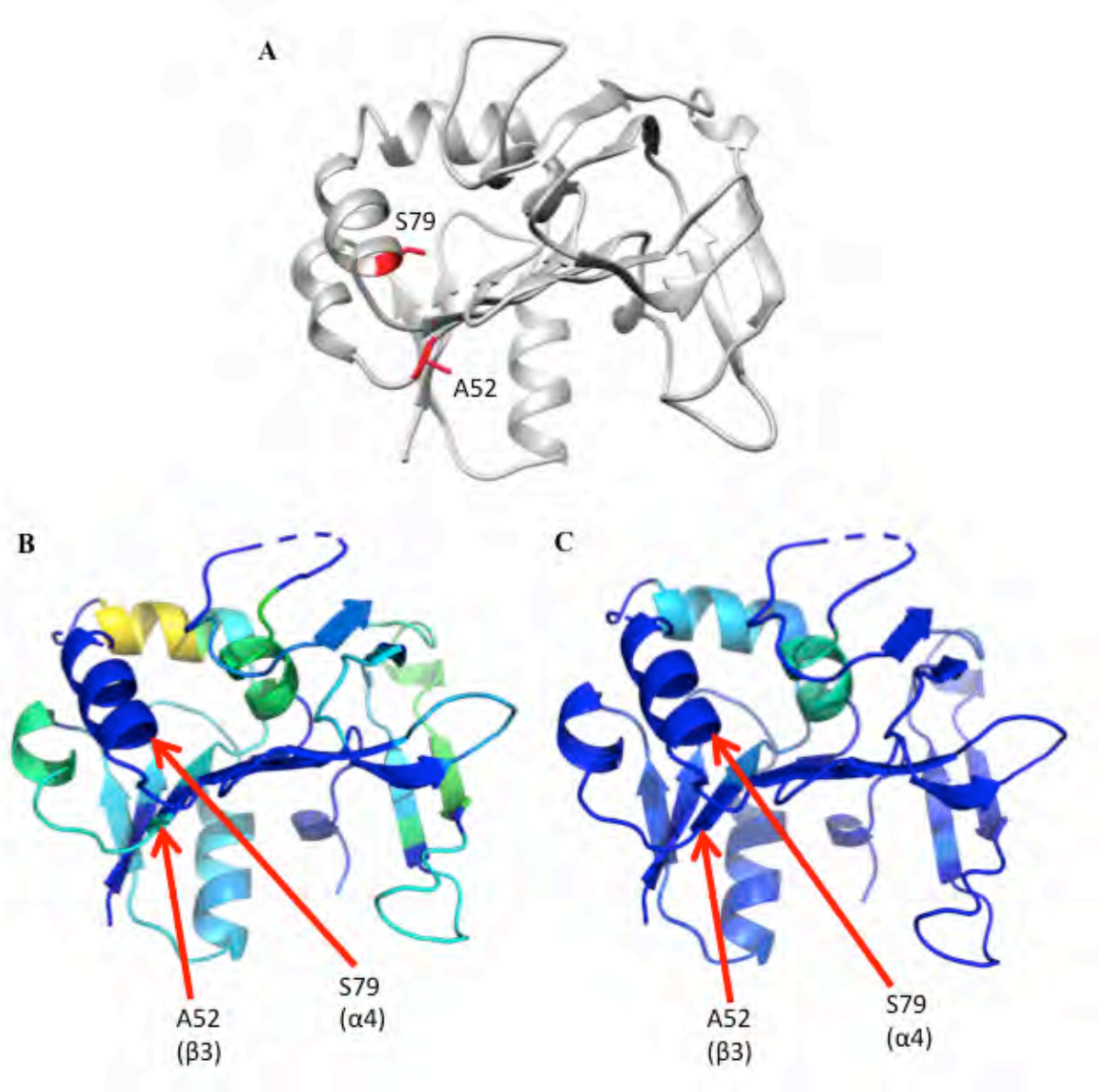


Figure 9. A) Mapping the *Sc*BPL residues analogous to the human BPL N-terminal extension mutations that cause Multiple Carboxylase Disease (MCD) (highlighted in red). Also shown, in the same structural orientation, are B) the apo-*Sc*BPL HDX results and C) the HDX difference results between apo-*Sc*BPL and holo-*Sc*BPL for the N-terminal domain.

References:

1. Tong, L. (2013) Structure and function of biotin-dependent carboxylases. *Cell Mol Life Sci* **70**, 863-891
2. Wolf, B. (2012) Biotinidase deficiency: "if you have to have an inherited metabolic disease, this is the one to have". *Genet Med* **14**, 565-575
3. Donti, T. R., Blackburn, P. R., and Atwal, P. S. (2016) Holocarboxylase synthetase deficiency pre and post newborn screening. *Mol Genet Metab Rep* **7**, 40-44
4. Baumgartner, M. R. (2013) Vitamin-responsive disorders: cobalamin, folate, biotin, vitamins B1 and E. *Handb Clin Neurol* **113**, 1799-1810
5. Chapman-Smith, A., and Cronan Jr, J. E. (1999) The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* **24**, 359-363
6. Bagautdinov, B., Kuroishi, C., Sugahara, M., and Kunishima, N. (2005) Crystal Structures of Biotin Protein Ligase from *Pyrococcus horikoshii* OT3 and its Complexes: Structural Basis of Biotin Activation. *J Mol Biol* **353**, 322-333
7. Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2008) Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate. *J Biol Chem* **283**, 14739-14750
8. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Pendini, N. R., Polyak, S. W., Pedersen, D. S., Morona, R., Turnidge, J. D., Wallace, J. C., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2012) Selective inhibition of Biotin Protein Ligase from *Staphylococcus aureus*. *J Biol Chem* **287**, 17823-17832
9. Ma, Q., Akhter, Y., Wilmanns, M., and Ehebauer, M. T. (2014) Active site conformational changes upon reaction intermediate biotinyl-5′-AMP binding in biotin protein ligase from *Mycobacterium tuberculosis*. *Protein Sci* **23**, 932-939
10. Purushothaman, S., Gupta, G., Srivastava, R., Ramu, V. G., and Surolia, A. (2008) Ligand Specificity of Group 1 Biotin Protein Ligase of *Mycobacterium tuberculosis*. *PLoS ONE* **3**
11. Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D., and Matthews, B. W. (2006) Co-repressor Induced Order and Biotin Repressor Dimerization: A Case for Divergent Followed by Convergent Evolution. *J Mol Biol* **357**, 509-523
12. Mayende, L., Swift, R. D., Bailey, L. M., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., and Polyak, S. W. (2012) A novel molecular mechanism to explain biotin-unresponsive holocarboxylase synthetase deficiency. *J Mol Med* **90**, 81-88
13. Weaver, L. H., Kwon, K., Beckett, D., and Matthews, B. W. (2001) Corepressor-induced organization and assembly of the biotin repressor: A model for allosteric activation of a transcriptional regulator. *PNAS* **98**, 6045-6050
14. Naganathan, S., and Beckett, D. (2007) Nucleation of an Allosteric Response *via* Ligand-induced Loop Folding. *J Mol Biol* **373**, 96-111
15. Gupta, V., Gupta, R. K., Khare, G., Salunke, D. M., Surolia, A., and Tyagi, A. K. (2010) Structural Ordering of Disordered Ligand-Binding Loops of Biotin Protein Ligase into Active Conformations as a Consequence of Dehydration. *PLoS ONE* **5**
16. Tron, C. M., McNae, I. W., Nutley, M., Clarke, D. J., Cooper, A., Walkinshaw, M. D., Baxter, R. L., and Campopiano, D. J. (2009) Structural and Functional Studies of the Biotin Protein Ligase from *Aquifex aeolicus* Reveal a Critical Role for a Conserved Residue in Target Specificity *J Mol Biol* **387**, 129-146
17. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterisation of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci* **22**, 762-773

18. Satiaputra, J., Shearwin, K. E., Booker, G. W., and Polyak, S. W. (2016) Mechanisms of biotin-regulated gene expression in microbes. *Synth Syst Biotechnol* **1**, 17-24
19. Polyak, S. W., Chapman-Smith, A., Brautigan, P. J., and Wallace, J. C. (1999) Biotin Protein Ligase from *Saccharomyces cerevisiae*. The N-terminal domain is required for complete activity. *J Biol Chem* **274**, 32847-32854
20. Campeau, E., and Gravel, R. A. (2001) Expression in *Escherichia coli* of N- and C-terminally Deleted Human Holocarboxylase Synthetase: Influence of the N-terminus on biotinylation and identification of a minimum functional protein. *J Biol Chem* **276**, 12310-12316
21. Hassan, Y. I., Moriyama, H., Olsen, L. J., Bi, X., and Zempleni, J. (2009) N- and C-terminal domains in human holocarboxylase synthetase participate in substrate recognition. *Mol Genet Metab* **96**, 183-188
22. Lee, C.-K., Cheong, C., and Jeon, Y. H. (2010) The N-terminal domain of human holocarboxylase synthetase facilitates biotinylation via direct interaction with the substrate protein. *FEBS Lett* **584**, 675-680
23. Sternicki, L. M., Wegener, K. L., Bruning, J. B., Booker, G. W., and Polyak, S. W. (2017) Mechanisms Governing Precise Protein Biotinylation. *Trends Biochem Sci* **42**, 383-394
24. Healy, S., McDonald, M. K., Wu, X., Yue, W. W., Kochan, G., Oppermann, U., and Gravel, R. A. (2010) Structural Impact of Human and *Escherichia coli* Biotin Carboxyl Carrier Proteins on Biotin Attachment. *Biochemistry* **49**, 4687-4694
25. Ingaramo, M., and Beckett, D. (2011) Biotinylation, a Post-translational Modification Controlled by the Rate of Protein-Protein Association. *J Biol Chem* **286**, 13071-13078
26. Ingaramo, M., and Beckett, D. (2012) Selectivity in Post-translational Biotin Addition to Five Human Carboxylases. *J Biol Chem* **287**, 1813-1822
27. Ingaramo, M., and Beckett, D. (2009) Distinct Amino Termini of Two Human HCS Isoforms Influence Biotin Acceptor Substrate Recognition. *J Biol Chem* **284**, 30862-30870
28. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C., Wallace, J. C., and Polyak, S. W. (2008) Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency. *Biochim Biophys Acta* **1784**, 973-982
29. Morrone, A., Malvagia, S., Donati, M. A., Funghini, S., Ciani, F., Pela, I., Boneh, A., Peters, H., Pasquini, E., and Zammarchi, E. (2002) Clinical findings and biochemical and molecular analysis of four patients with holocarboxylase synthetase deficiency. *Am J Med Genet* **111**, 10-18
30. Dupuis, L., Leon-Del-Rio, A., Leclerc, D., Campeau, E., Sweetman, L., Saudubray, J. M., Herman, G., Gibson, K. M., and Gravel, R. A. (1996) Clustering of mutations in the biotin-binding region of holocarboxylase synthetase in biotin-responsive multiple carboxylase deficiency. *Hum Mol Genet* **5**, 1011-1016
31. Wilson, C. J., Myer, M., Darlow, B. A., Stanley, T., Thomson, G., Baumgartner, E. R., Kirby, D. M., and Thorburn, D. R. (2005) Severe Holocarboxylase Synthetase Deficiency with Incomplete Biotin Responsiveness Resulting in Antenatal Insult in Samoan Neonates. *J Pediatr* **147**, 115-118
32. Esaki, S., Malkaram, S. A., and Zempleni, J. (2012) Effects of single-nucleotide polymorphisms in the human holocarboxylase synthetase gene on enzyme catalysis. *Eur J Hum Genet* **20**, 428-433
33. Dupuis, L., Campeau, E., Leclerc, D., and Gravel, R. A. (1999) Mechanism of Biotin Responsiveness in Biotin-Responsive Multiple Carboxylase Deficiency. *Mol Genet Metab* **66**, 80-90
34. Sakamoto, O., Suzuki, Y., Li, X., Aoki, Y., Hiratsuka, M., Suormala, T., Baumgartner, E. R., Gibson, K. M., and Narisawa, K. (1999) Relationship between kinetic properties

of mutant enzyme and biochemical and clinical responsiveness to biotin in holocarboxylase synthetase deficiency. *Pediatr Res* **46**, 671-676
35. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C. J., Wallace, J. C., and Polyak, S. W. (2008) Biotin protein ligase from *Candida albicans*: Expression, purification and development of a novel assay. *Arch Biochem Biophys* **479**, 163-169
36. Satiaputra, J., Sternicki, L. M., Hayes, A. J., Pukala, T. L., Booker, G. W., Shearwin, K. E., and Polyak, S. W. (2019) Native mass spectrometry identifies an alternate DNA-binding pathway for BirA from *Staphylococcus aureus*. *Scientific Reports* **9**
37. Soares da Costa, T. P., Yap, M. Y., Perugini, M. A., Wallace, J. C., Abell, A. D., Wilce, M. C. J., Polyak, S. W., and Booker, G. W. (2014) Dual roles of F123 in protein homodimerization and inhibitor binding to biotin protein ligase from *Staphylococcus aureus*. *Mol Microbiol* **91**, 110-120
38. Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V., and Ruotolo, B. T. (2010) Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal Chem* **82**, 9557-9565
39. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc* **3**, 1139-1152
40. Migas, L. G., France, A. P., Bellina, B., and Barran, P. E. (2017) ORIGAMI: A software suite for activated ion mobility mass spectrometry (aIM-MS) applied to multimeric protein assemblies. *Int J Mass Spectrom* **427**, 20-28
41. Boivin, S., Kozak, S., and Meijers, R. (2013) Optimization of protein purification and characterization using Thermofluor screens. *Protein Expr Purif* **91**, 192-206
42. Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J., and Weiss, M. S. (2013) Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**, 209-214
43. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858
44. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303
45. Chen, V. B., Arendall III, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21
46. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall III, W. B., Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-W383
47. Govindaraju, V., Basus, V. J., Matson, G. B., and Maudsley, A. A. (1998) Measurement of Chemical Shifts and Coupling Constants for Glutamate and Gutamine. *Magn Reson Med* **39**, 1011-1013
48. Satiaputra, J., Sternicki, L. M., Hayes, A. J., Pukala, T. L., Booker, G. W., Shearwin, K. E., and Polyak, S. W. (2019) Native mass spectrometry identifies an alternate DNA-binding pathway for BirA from *Staphylococcus aureus*. *Sci Rep* **9**
49. Pukala, T. (2018) Importance of Collision Cross Section Measurements by Ion Mobility-Mass Spectrometry in Structural Biology. *Rapid Commun Mass Spectrom*
50. Dixit, S. M., Polasky, D. A., and Ruotolo, B. T. (2018) Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Curr Opin Chem Biol* **42**, 93-100

51. Hopper, J. T. S., and Oldham, N. J. (2009) Collision Induced Unfolding of Protein Ions in the Gas Phase Studied by Ion Mobility-Mass Spectrometry: The Effect of Ligand Binding on Conformational Stability. *J Am Soc Mass Spectrom* **20**, 1851-1858
52. Niu, S., and Ruotolo, B. T. (2015) Collisional unfolding of multiprotein complexes reveals cooperative stabilization upon ligand binding. *Protein Sci* **24**, 1272-1281
53. Lee, J.-J., Park, Y. S., and Lee, K.-J. (2015) Hydrogen-deuterium exchange mass spectrometry for determining protein structural changes in drug discovery. *Arch Pharm Res* **38**, 1737-1745
54. Jaswal, S. S. (2013) Biological insights from hydrogen exchange mass spectrometry. *Biochim Biophys Acta* **1834**, 1188-1201
55. Yan, X., Watson, J., Ho, P. S., and Deinzer, M. L. (2004) Mass spectrometric approaches using electrospray ionization charge states and hydrogen-deuterium exchange for determining protein structures and their conformational changes. *Mol Cell Proteomics* **3**, 10-23
56. Reche, P. A. (2000) Lipoylating and biotinylating enzymes contain a homologous catalytic module. *Protein Sci* **9**, 1922-1929
57. Zein, F., Zhang, Y., Kang, Y.-N., Burns, K., Begley, T. P., and Ealick, S. E. (2006) Structural Insights into the Mechanism of the PLP Synthase Holoenzyme from *Thermotoga maritima Biochemistry* **45**, 14609-14620
58. Korolev, S., Skarina, T., Evdokimova, E., Beasley, S., Edwards, A., Joachimiak, A., and Savchenko, A. (2002) Crystal Structure of Glutamine Amidotransferase from *Thermotoga maritima*. *Proteins* **49**, 420-422
59. Mouilleron, S., and Golinelli-Pimpaneau, B. (2007) Conformational changes in ammonia-channeling glutamine amidotransferases. *Curr Opin Struct Biol* **17**, 653-664
60. van den Heuvel, R. H. H., Curti, B., Vanoni, M. A., and Mattevi, A. (2004) Glutamate synthase: a fascinating pathway from L-glutamine to L-glutamate. *Cell Mol Life Sci* **61**, 669-681
61. Aoki, Y., Suzuki, Y., Sakamoto, O., Li, X., Takahashi, K., Ohtake, A., Sakuta, R., Ohura, T., Miyabayashi, S., and Narisawa, K. (1995) Molecular analysis of holocarboxylase synthetase deficiency: a missense mutation and a single base deletion are predominant in Japanese patients. *Biochim Biophys Acta* **1272**, 168-174
62. Suzuki, Y., Yang, X., Aoki, Y., Kure, S., and Matsubara, Y. (2005) Mutations in the Holocarboxylase Synthetase Gene *HLCS*. *Hum Mutat* **26**, 285-290

SUPPLEMENTARY INFORMATION

Structural investigations of the N-terminal domain of the class III *Saccharomyces cerevisiae* Biotin Protein Ligase

Louise M Sternicki[1]*, Tara L Pukala[2], Kamila Pacholarz[3], Perdita Barran[3], Grant W Booker[1], Steven W Polyak[1, 4, 5 +] and Kate L Wegener[1, 4 *+]

[1] School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

[2] School of Physical Sciences, The University of Adelaide, South Australia 5005, Australia

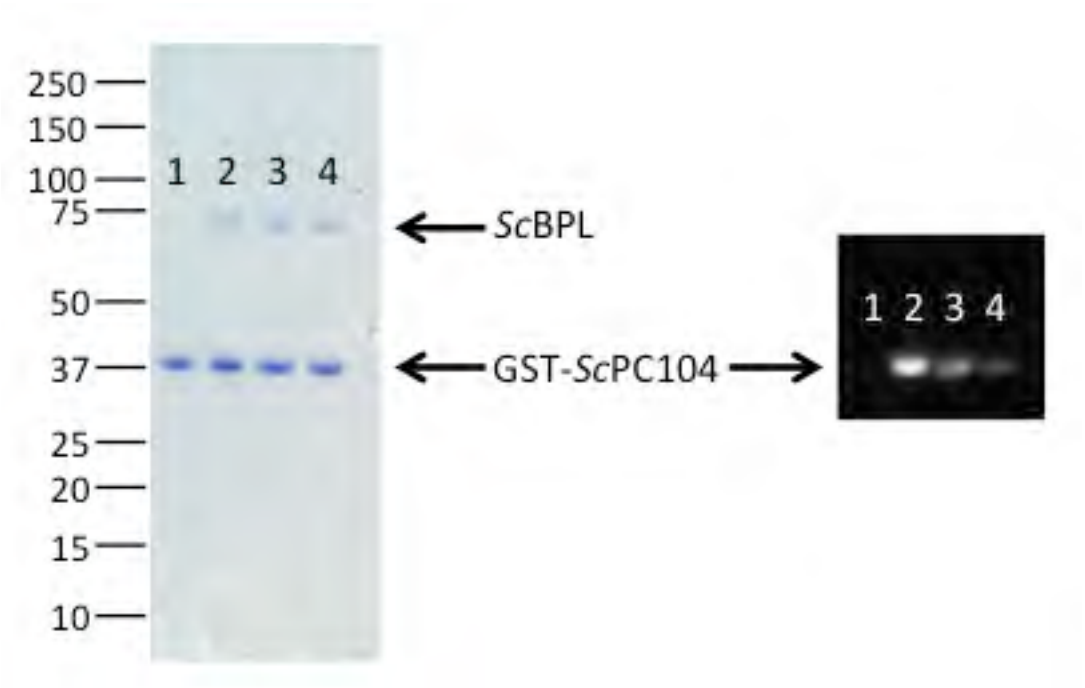
[3] Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, United Kingdom

[4] Institute for Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, South Australia 5005, Australia

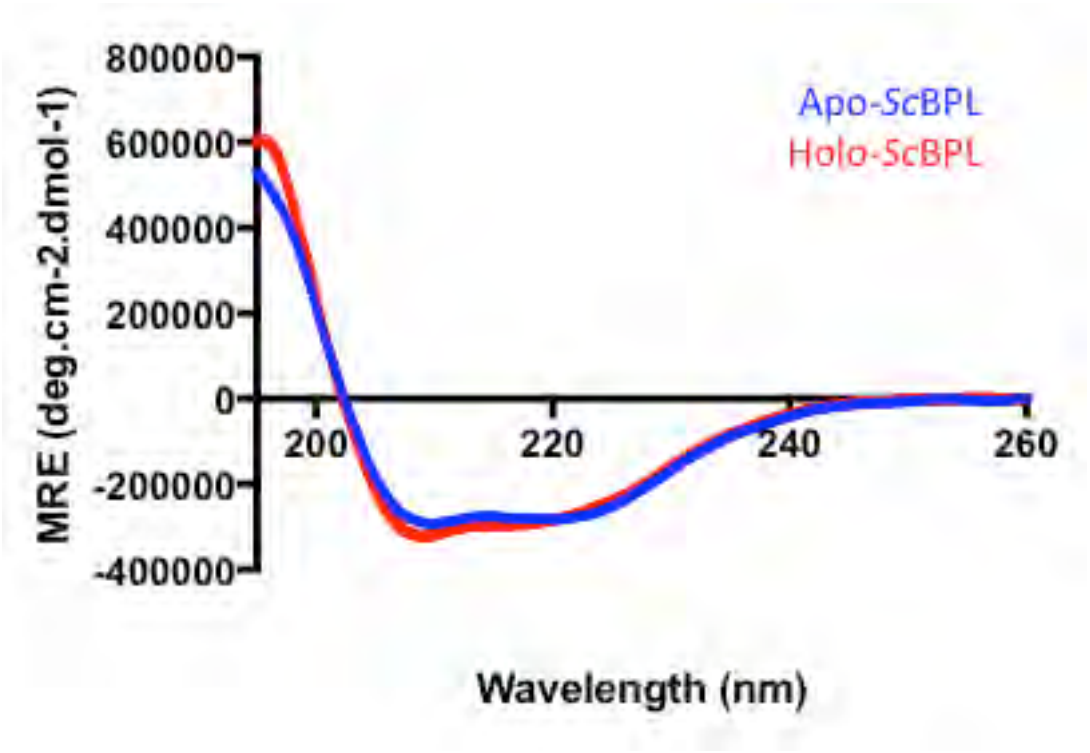
[5] Current address: School of Pharmacy and Medical Sciences, University of South Australia, 5001, Australia

* To whom correspondence should be addressed.

[+] Equal senior authors



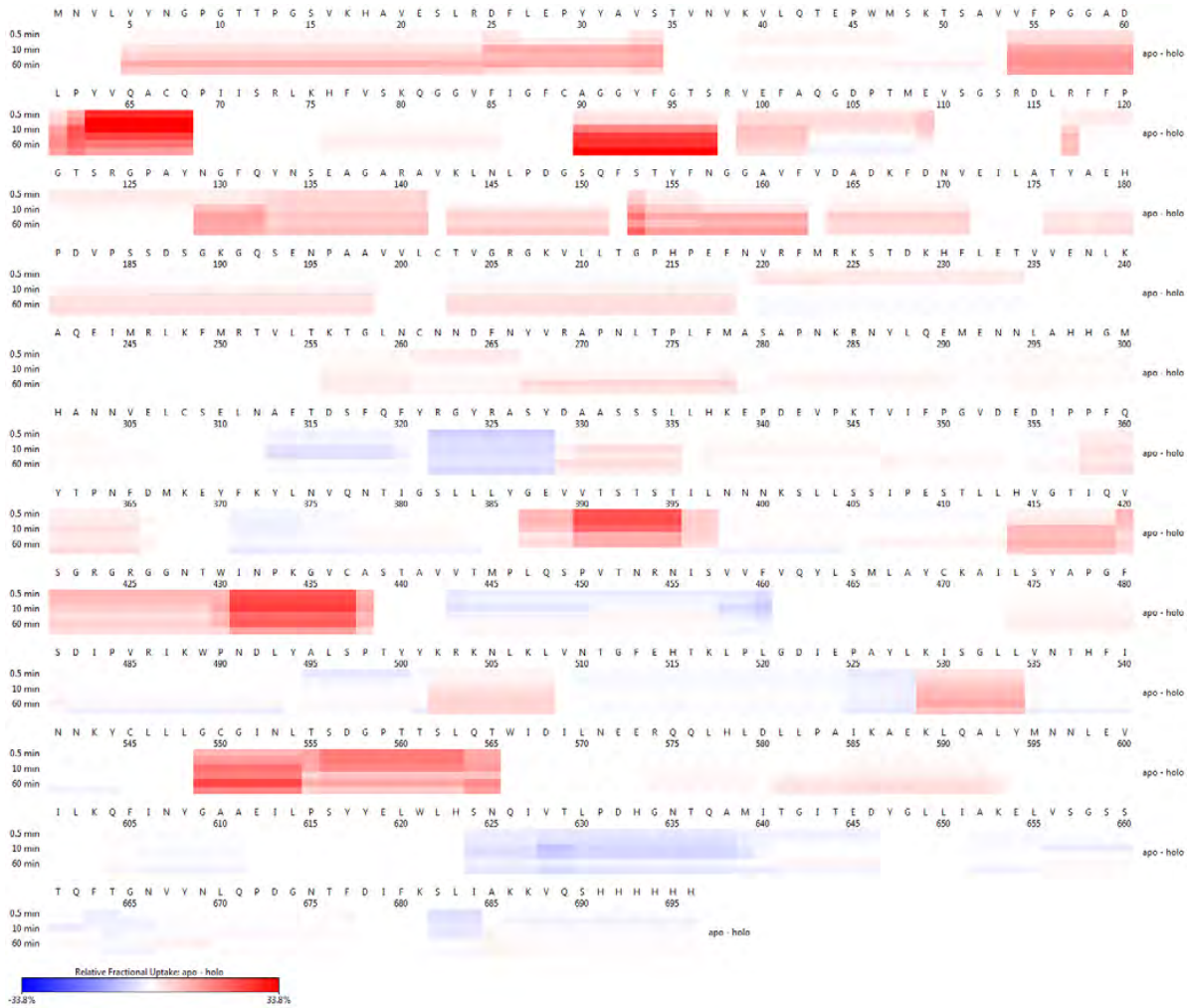
Supplementary Figure 1. Biotinylation blot confirms the production of apo-*Sc*BPL. Apo-*Sc*BPL was measured for its ability to biotinylate the GST-conjugated biotin domain GST-*Sc*PC104 – the biotin domain from the *S. cerevisiae* pyruvate carboxylase isoform 1. Reactions containing 50 mM Tris pH 8.0, 3 mM ATP, 5.5 mM $MgCl_2$, 0.1 mM DTT, 15 µM GST-*Sc*PC104 and 3 µM *Sc*BPL were incubated in the presence and absence of biotin for 1 hour at 37 °C. Biotinylation was detected utilising a streptavidin-conjugated fluorophore (Alexa488). Biotinylation of GST-*Sc*PC104 by apo-*Sc*BPL in the absence of biotin indicates remaining bound biotin or biotinyl-5′-AMP from expression and purification. Lanes include 1) 15 µM apo-GST-*Sc*PC104 to measure the intrinsic biotinylation of the GST-*Sc*PC104 protein preparation (no *Sc*BPL or biotin in reaction), 2) 15 µM holo-GST-*Sc*PC104 to act as a measure of complete GST-*Sc*PC104 biotinylation (15 µM biotin in reaction), 3) 3 µM holo-GST-*Sc*PC104 to measure the biotinylation signal from 3 µM of holo-*Sc*BPL (3 µM biotin in reaction), and 4) 3 µM apo-*Sc*BPL + 15 µM apo-GST-*Sc*PC104 to measure the biotinylation from the apo-*Sc*BPL preparation. A faint band present in lane 4 indicates there was some biotin or biotinyl-5-AMP in the apo-*Sc*BPL preparation that could biotinylate the apo-GST-*Sc*PC104 (measured to have no intrinsic biotinylation by lane 1).



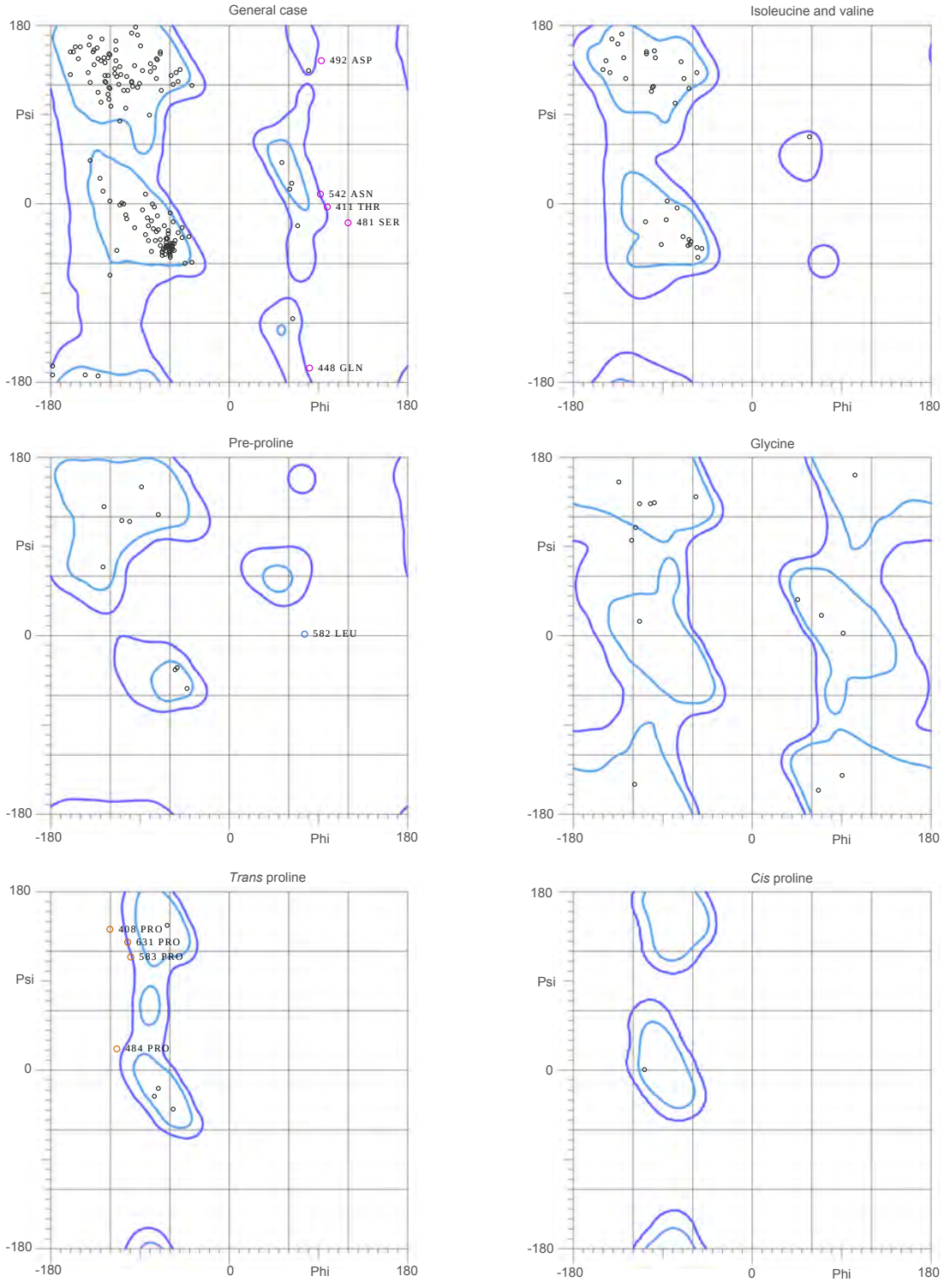
Supplementary Figure 2. There is no difference in the secondary structure of apo- (blue) and holo-*Sc*BPL (red), as measured by CD spectroscopy. Spectra are the smoothed average of 5 scans corrected for buffer and protein concentration.

Supplementary Figure 3. Sequence coverage of *Sc*BPL by the peptides identified from HDX LC-MS. A total of 149 peptides were identified, resulting in 85.5% coverage of *Sc*BPL with a redundancy of 2.99.

Supplementary Figure 4. The HDX results for apo- and holo-*Sc*BPL mapped across the sequence of *Sc*BPL. Yellow colouring signifies the incorporation of deuterium. The linker region (resides 220-380), which was not modelled, is shown in the red box.



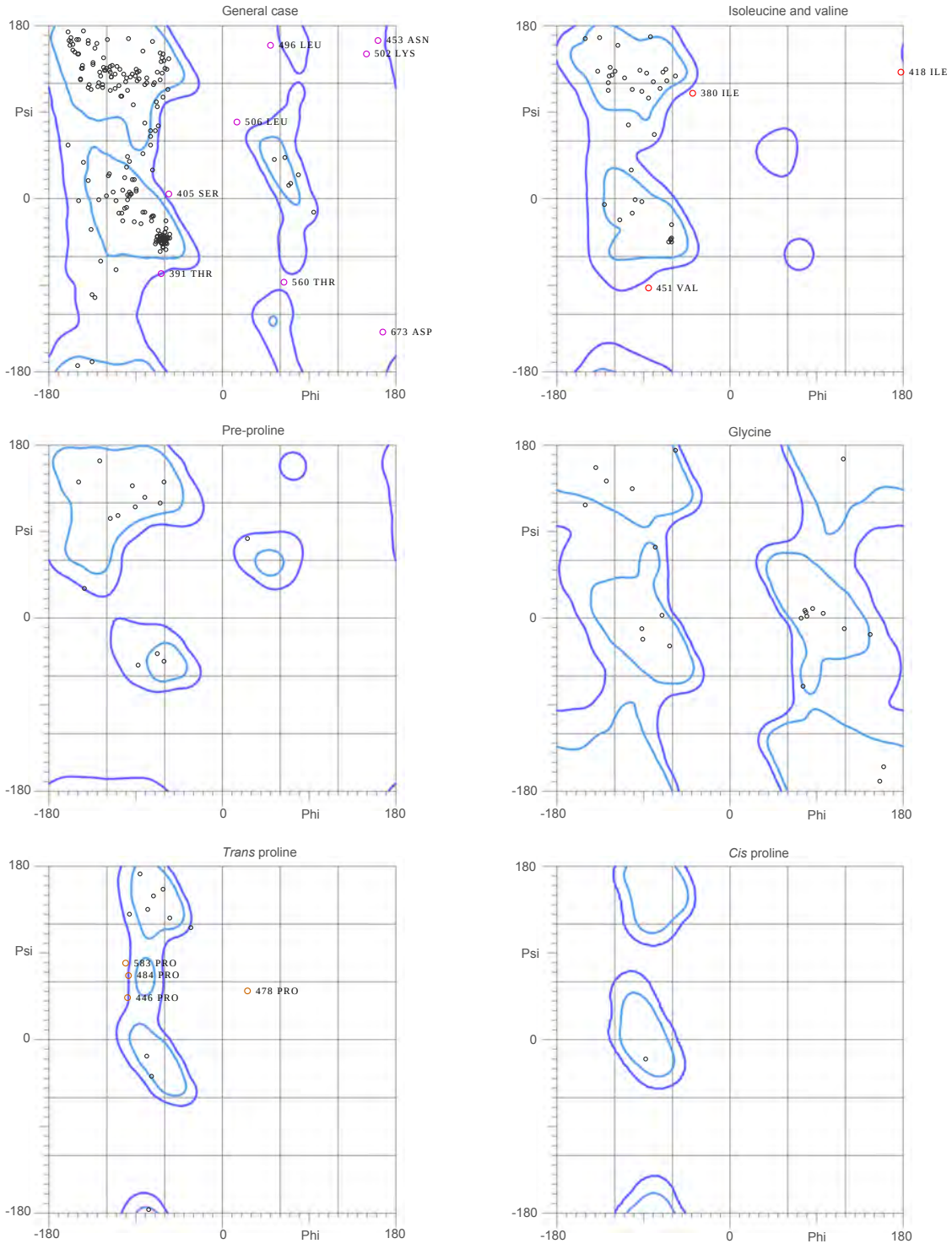
Supplementary Figure 5. Heat map identifying the difference in deuterium incorporation between apo- and holo-*Sc*BPL across the sequence. Red colouring indicates greater deuterium incorporation for apo-*Sc*BPL, whilst blue indicates more deuterium uptake in holo-*Sc*BPL.



89.1% (205/230) of all residues were in favored (98%) regions.
95.7% (220/230) of all residues were in allowed (>99.8%) regions.

There were 10 outliers (phi, psi):
408 PRO (-121.8, 143.9)
411 THR (99.3, -4.0)
448 GLN (81.7, -166.9)
481 SER (120.4, -19.9)
484 PRO (-114.5, 22.6)
492 ASP (93.4, 145.6)
542 ASN (92.6, 10.3)
582 LEU (76.9, 2.1)
583 PRO (-100.5, 115.2)
631 PRO (-103.7, 130.3)

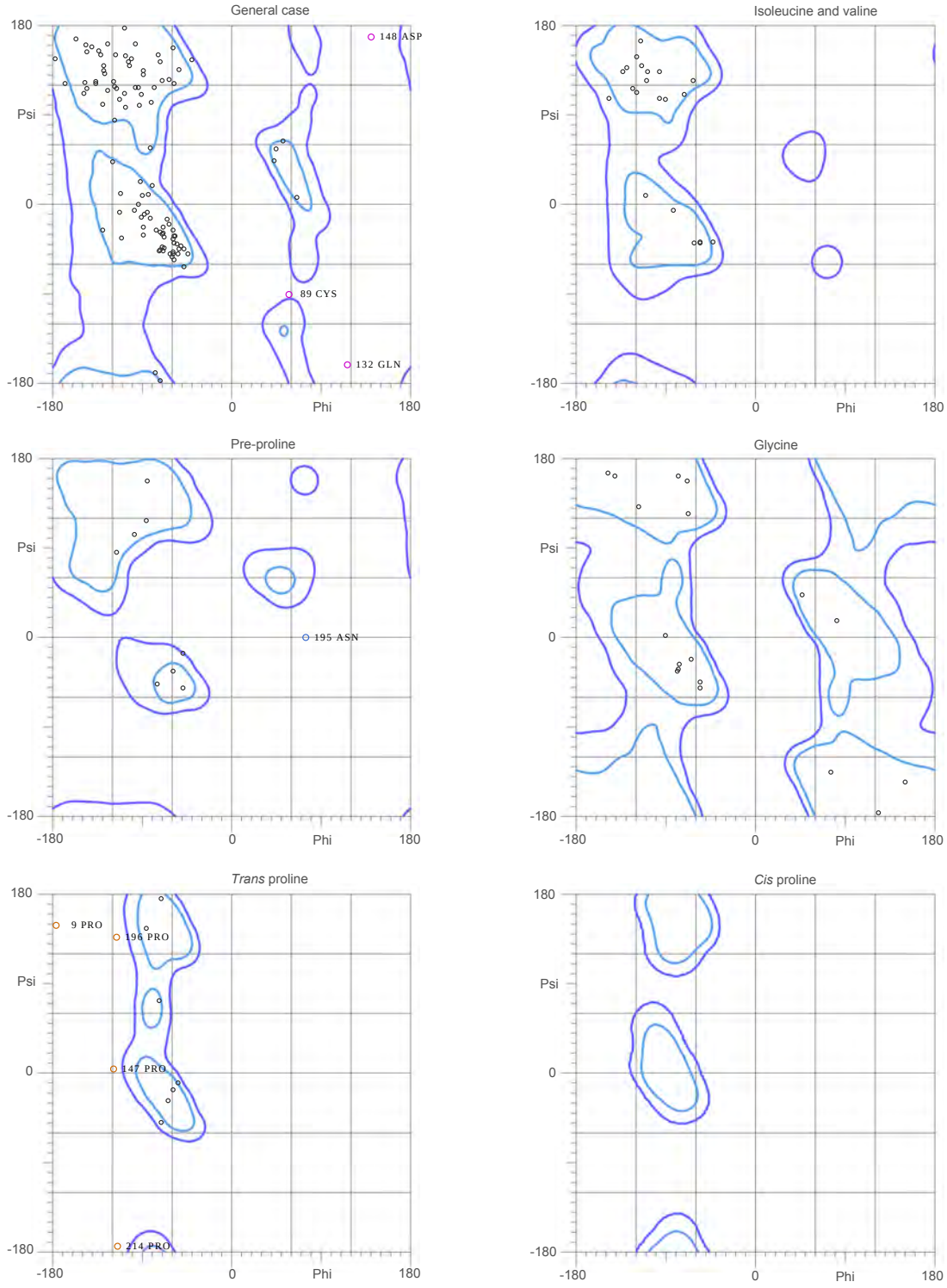
Supplementary Figure 6. Ramachandran plots of the residues in the *Sc*BPL catalytic domain homology model generated by Phyre (1). Plots were produced using MolProbity (2,3).



83.2% (253/304) of all residues were in favored (98%) regions.
95.1% (289/304) of all residues were in allowed (>99.8%) regions.

There were 15 outliers (phi, psi):
380 ILE (-39.4, 110.5)
391 THR (-64.1, -78.6)
405 SER (-56.4, 5.1)
418 ILE (178.6, 132.7)
446 PRO (-99.1, 44.4)
451 VAL (-86.0, -93.9)
453 ASN (162.4, 165.7)
478 PRO (26.4, 51.8)
484 PRO (-99.0, 67.9)
496 LEU (50.6, 160.1)
502 LYS (150.1, 151.4)
506 LEU (15.5, 80.6)
560 THR (64.6, -87.9)
583 PRO (-101.7, 81.0)
673 ASP (167.9, -139.1)

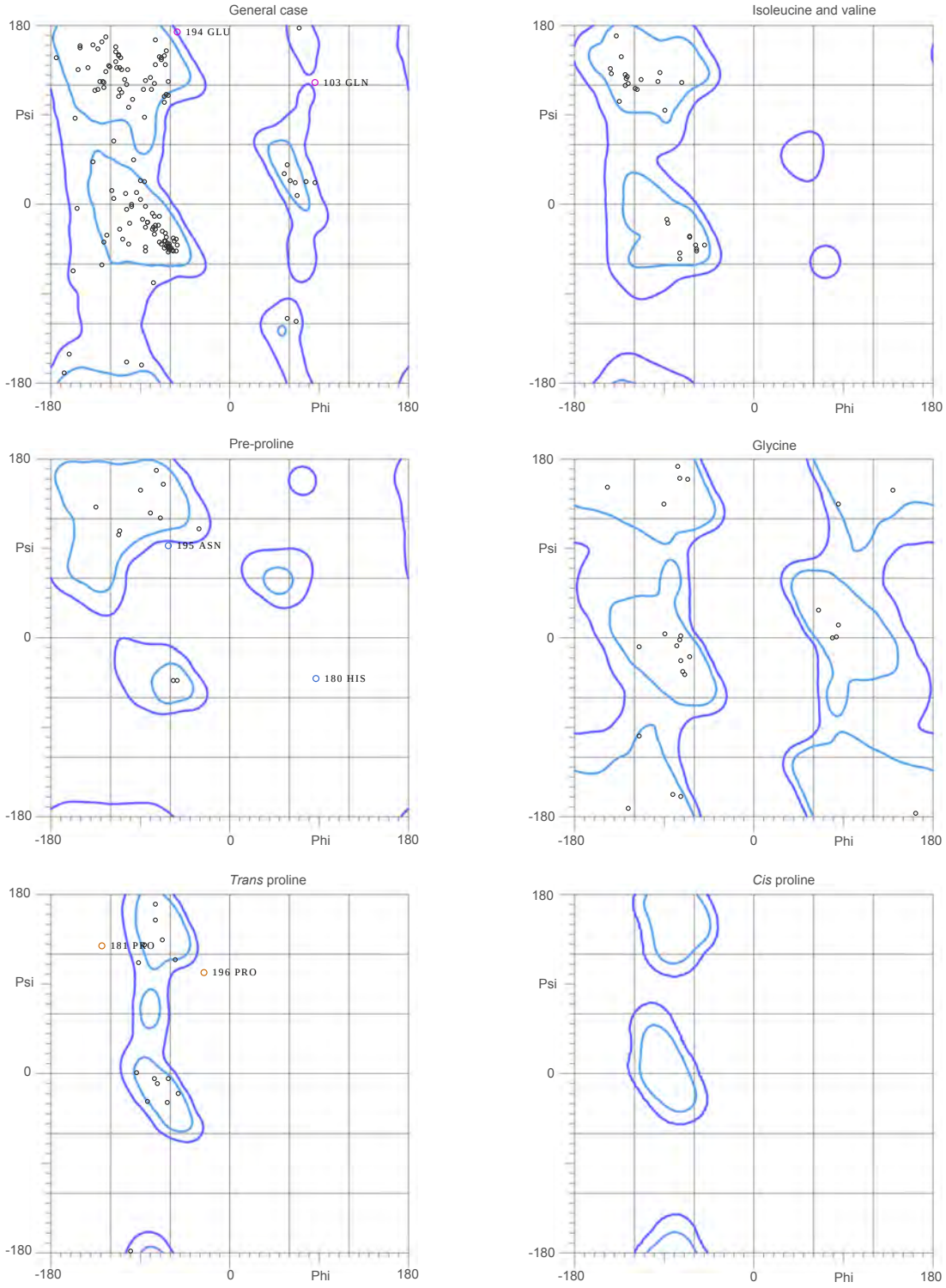
Supplementary Figure 7. Ramachandran plots of the residues in the *Sc*BPL catalytic domain homology model generated by SWISS-MODEL (4). Plots were produced using MolProbity (2,3).



88.5% (146/165) of all residues were in favored (98%) regions.
95.2% (157/165) of all residues were in allowed (>99.8%) regions.

There were 8 outliers (phi, psi):
9 PRO (-177.8, 149.7)
89 CYS (58.4, -91.1)
132 GLN (117.4, -162.9)
147 PRO (-119.6, 4.7)
148 ASP (141.1, 169.5)
195 ASN (75.1, -0.4)
196 PRO (-116.3, 138.0)
214 PRO (-115.3, -175.8)

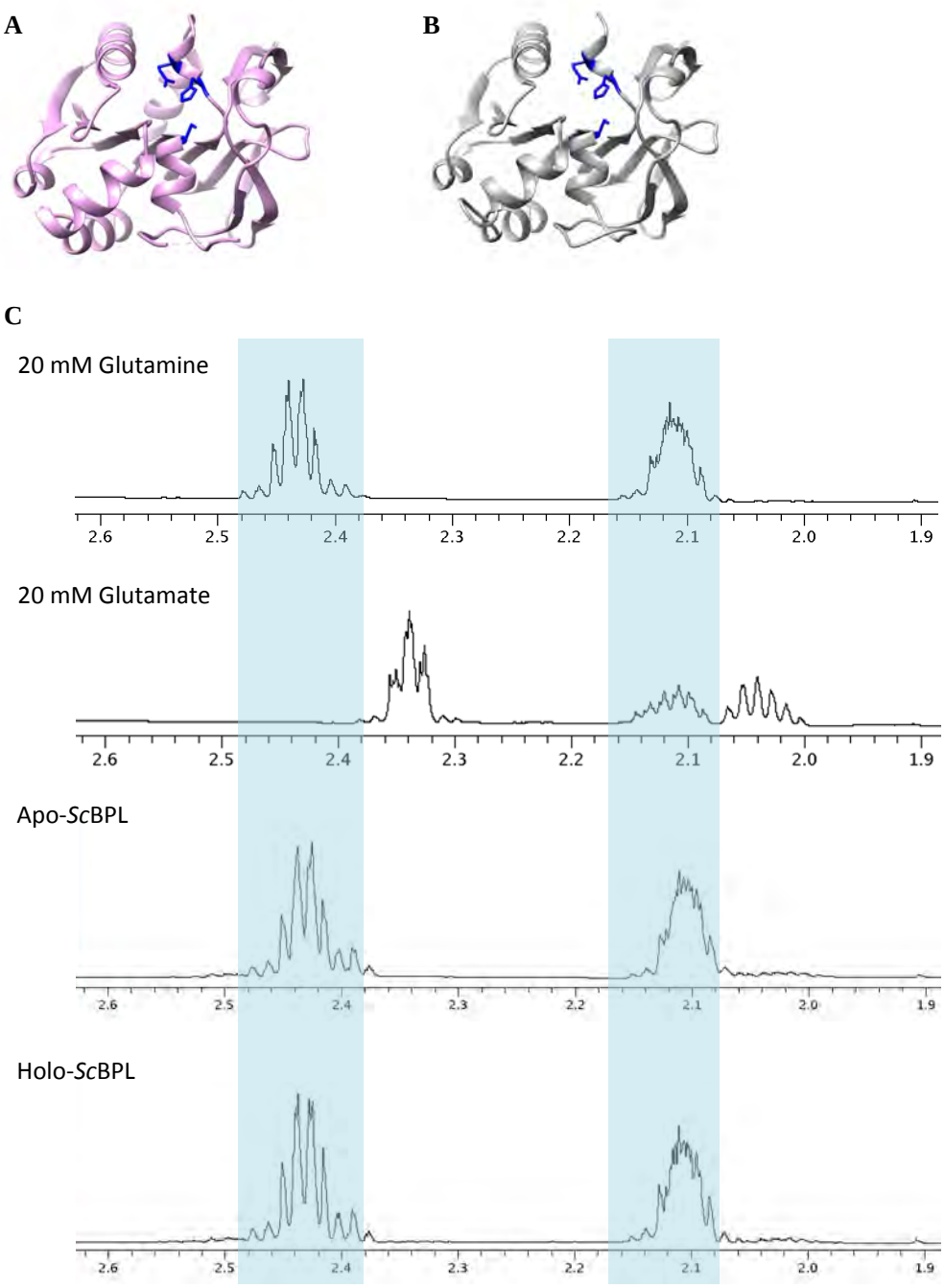
Supplementary Figure 8. Ramachandran plots of the residues in the *Sc*BPL N-terminal domain homology model generated by Phyre (1). Plots were produced using MolProbity (2,3).



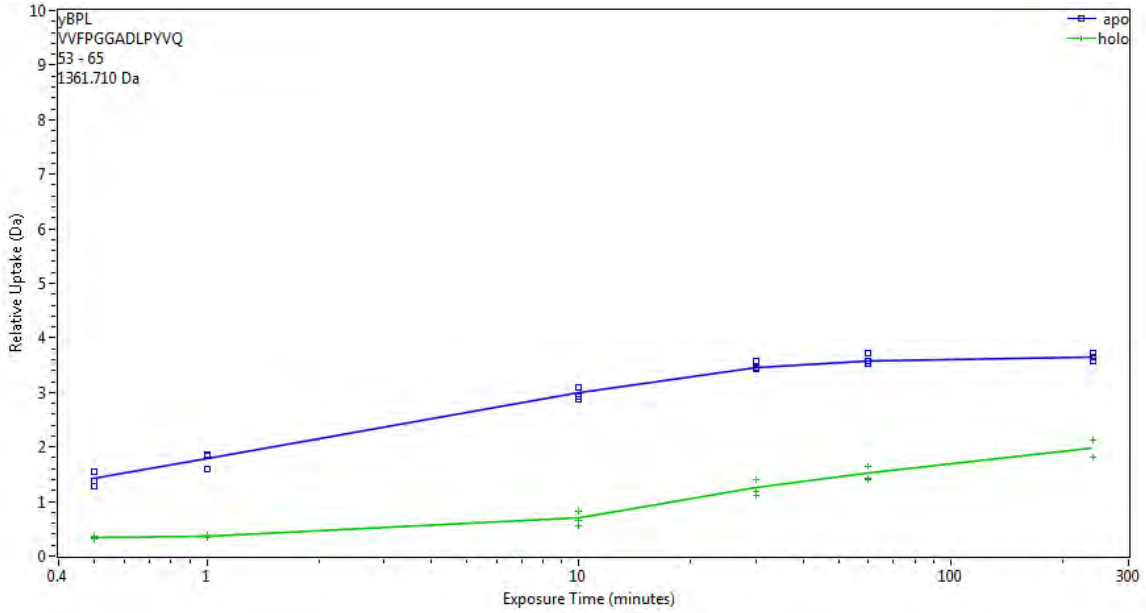
85.8% (187/218) of all residues were in favored (98%) regions.
97.2% (212/218) of all residues were in allowed (>99.8%) regions.

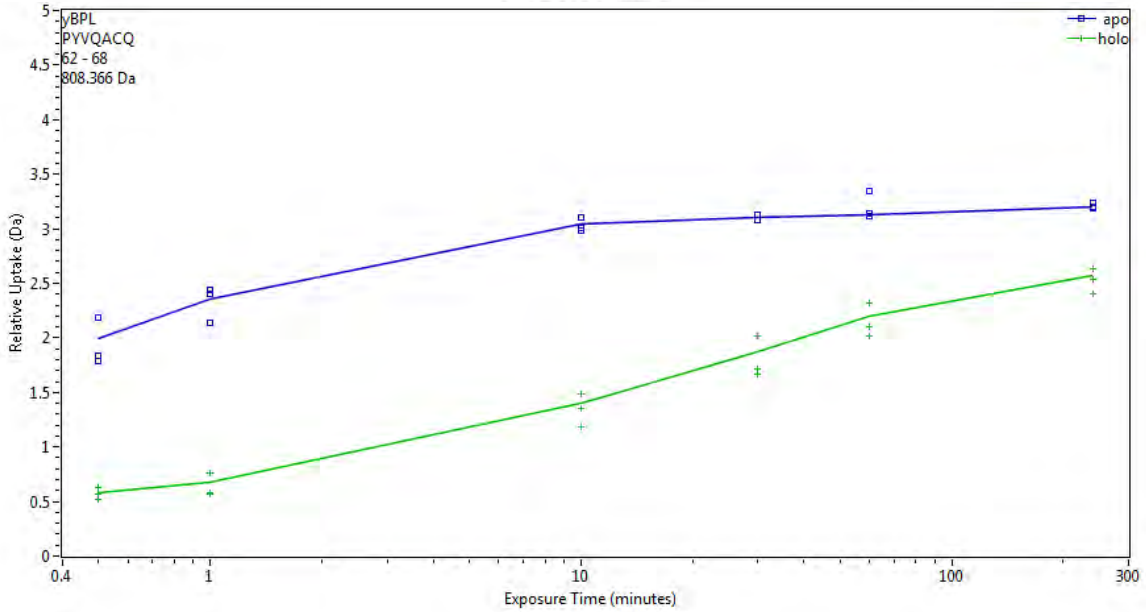
There were 6 outliers (phi, psi):
103 GLN (86.6, 123.4)
180 HIS (87.3, -41.4)
181 PRO (-130.0, 129.1)
194 GLU (-53.4, 174.5)
195 ASN (-62.2, 93.1)
196 PRO (-26.9, 102.3)

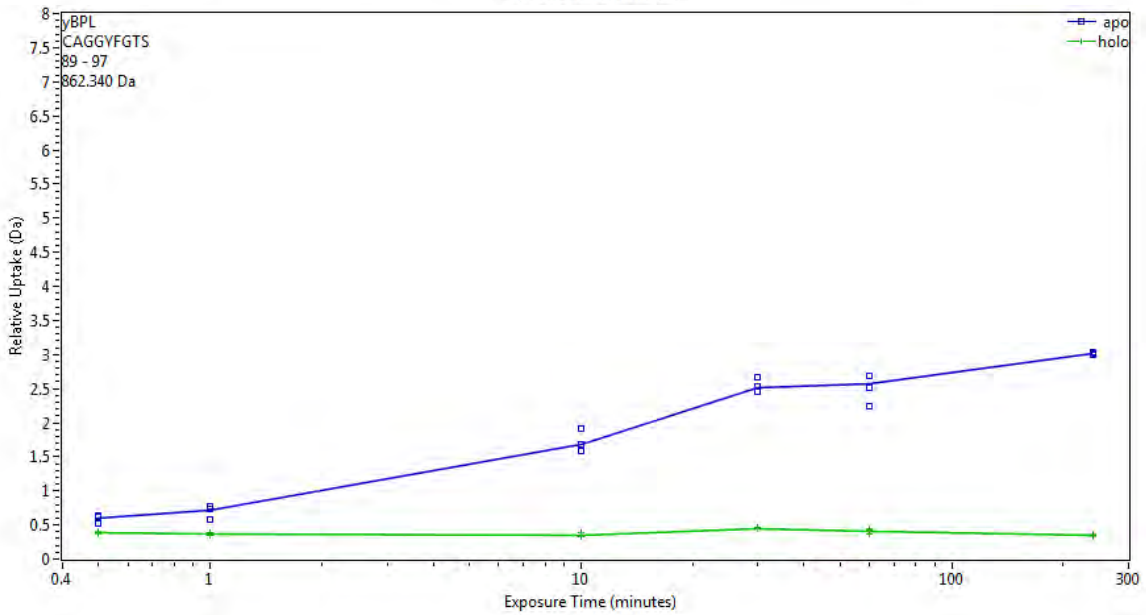
Supplementary Figure 9. Ramachandran plots of the residues in the *Sc*BPL N-terminal domain homology model generated by SWISS-MODEL (4). Plots were produced using MolProbity (2,3).

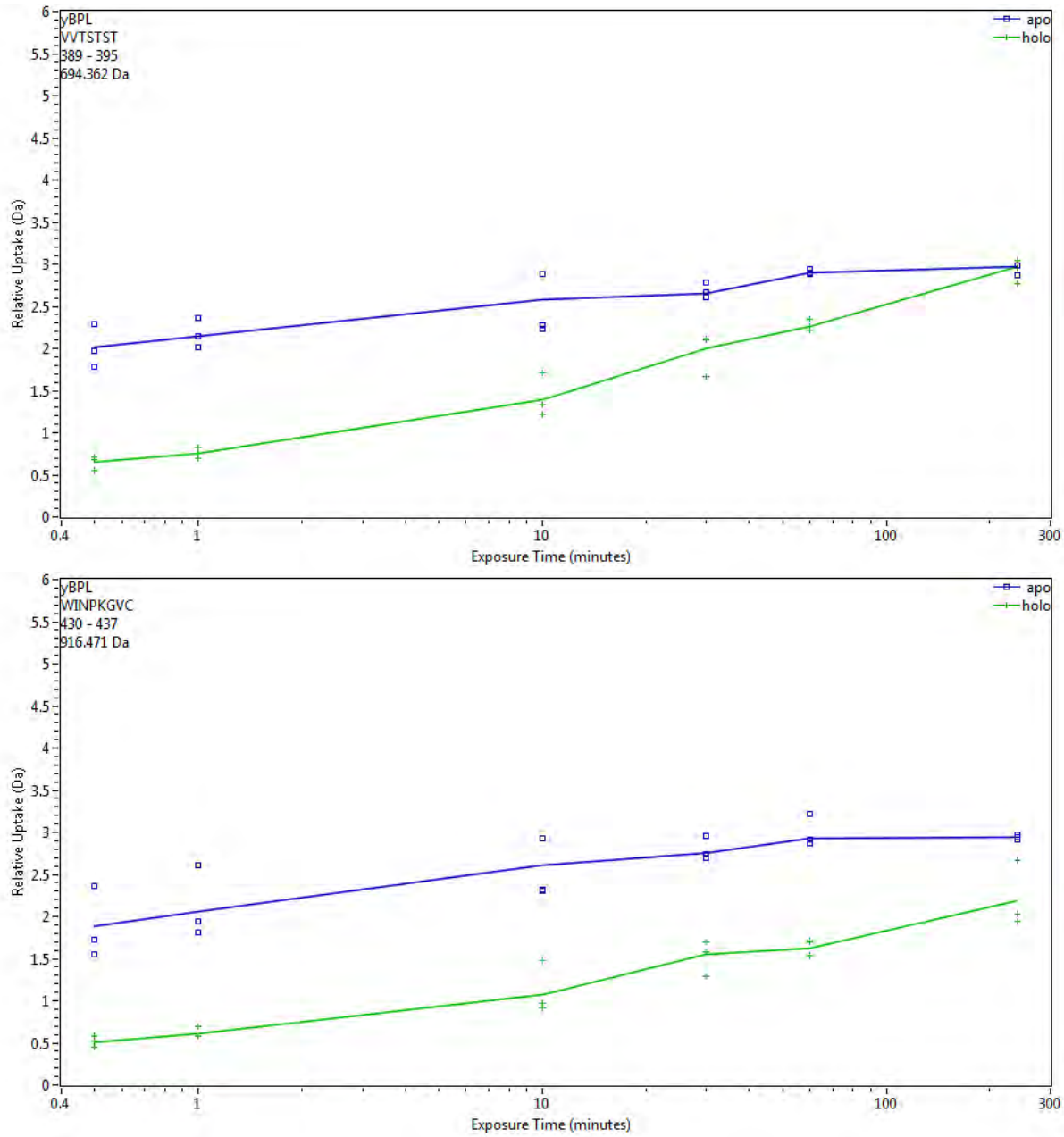


Supplementary Figure 10. *Sc*BPL does not contain glutamine amidotransferase (GATase) activity as measured by ^{1}H 1D NMR. A) GATases contain a catalytic triad of residues (cysteine, histidine and glutamic acid) (PDB: 2ISS, glutamine amidotransferase from the *Thermotoga maritima* pyridoxal 5′-phosphate holoenzyme (5)) B) that were conserved and correctly positioned in the Phyre homology model of the *Sc*BPL N-terminal domain (C89, H215, E217). C) Both apo- and holo-*Sc*BPL (biotin and MgATP) were incubated with glutamine and the reaction monitored over 4 days for conversion to glutamate. Controls of 20 mM glutamine and glutamate were used for chemical shift comparison.


yBPL
VVFPGGADLPYVQ
53 - 65
1361.710 Da
apo
holo
Relative Uptake (Da)
Exposure Time (minutes)


yBPL
PYVQACQ
62 - 68
808.366 Da
apo
holo
Relative Uptake (Da)
Exposure Time (minutes)


yBPL
CAGGYFGTS
89 - 97
862.340 Da
apo
holo
Relative Uptake (Da)
Exposure Time (minutes)



Supplementary Figure 11. Examples of the deuterium uptake rates for specific peptides measured over the four-hour time-course. Uptake rates for apo-*Sc*BPL shown in blue and rates for holo-*Sc*BPL in green. The corresponding residue numbering for the peptide position within *Sc*BPL is annotated.

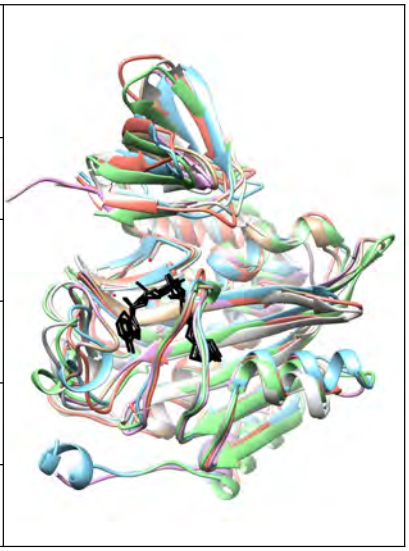
Supplementary Table 1. Masses of the species observed in the MS spectra of apo- and holo-treated *Sc*BPL.

***Sc*BPL**	**Measured MW (Da)**	**Corresponding ligand-bound state**	**Calculated Mass (Da)**
apo-	77261	apo, no ligands bound	77244
holo-	77275	apo, no ligands bound	77244
	77817	holo, reaction intermediate (biotinyl-5′-AMP) bound	77817

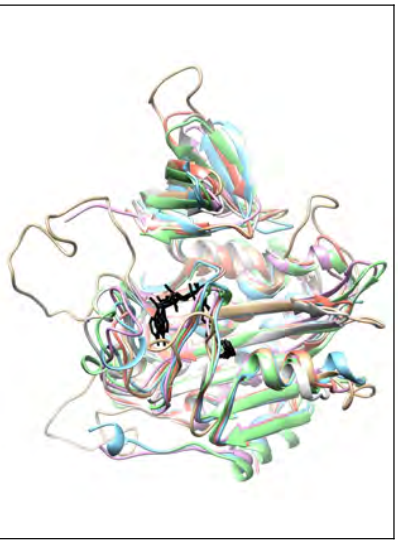
Supplementary Table 2. Mean, standard deviation (SD) and full width half maximum (FWHM) calculated from the IM-MS CCS distribution data for all charge states (15+ to 18+) for apo-*Sc*BPL and holo-*Sc*BPL.

		Apo-*Sc*BPL	**Holo-*Sc*BPL**
18+	MEAN (nm^2)	52.31	52.06
	SD (nm^2)	2.53	1.31
	FWHM (nm^2)	3.27	1.98
17+	MEAN (nm^2)	51.07	50.96
	SD (nm^2)	1.87	1.32
	FWHM (nm^2)	3.40	2.78
16+	MEAN (nm^2)	50.6	50.6
	SD (nm^2)	1.93	1.46
	FWHM (nm^2)	4.30	3.49
15+	MEAN (nm^2)	50.32	50.69
	SD (nm^2)	2.41	1.63
	FWHM (nm^2)	4.15	3.10

Supplementary Table 3. Structural similarity, measured by RMSD, between the *Sc*BPL C-terminal catalytic domain homology model produced using Phyre and crystal structures of the BPLs from other species. An overlay of the structures is also shown (model coloured beige). References for the structures are 4OP0 (6), 1WPY (7), 2EAY (8), 2EWN (9) and 3RIR (10).

Species	**BPL Class**	**PDB**	**Colour**	**RMSD Å2 between secondary structure cores (and whole length structures)**	
M. tuberculosis	Class I	4OP0	blue	1.720 (3.796)	
P. horikoshii	Class I	1WPY	red	1.427 (3.429)	
A. aquifex	Class I	2EAY	grey	1.373 (4.095)	
E. coli	Class II	2EWN	purple	0.001 (1.181)	
S. aureus	Class II	3RIR	green	1.526 (2.779)	

Supplementary Table 4. Structural similarity, measured by RMSD, between the *Sc*BPL C-terminal catalytic domain homology model produced using SWISS-MODEL and crystal structures of the BPLs from other species. An overlay of the structures is also shown (model coloured beige). References for the structures are 4OP0 (6), 1WPY (7), 2EAY (8), 2EWN (9) and 3RIR (10).

Species	**BPL Class**	**PDB**	**Colour**	**RMSD Å2 between secondary structure cores (and whole length structures)**	
M. tuberculosis	Class I	4OP0	blue	1.281 (7.828)	
P. horikoshii	Class I	1WPY	red	0.572 (5.896)	
A. aquifex	Class I	2EAY	grey	1.351 (8.520)	
E. coli	Class II	2EWN	purple	1.578 (9.134)	
S. aureus	Class II	3RIR	green	1.587 (7.532)	

References:

1. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858
2. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall III, W. B., Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-W383
3. Chen, V. B., Arendall III, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21
4. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303
5. Zein, F., Zhang, Y., Kang, Y.-N., Burns, K., Begley, T. P., and Ealick, S. E. (2006) Structural Insights into the Mechanism of the PLP Synthase Holoenzyme from *Thermotoga maritima Biochemistry* **45**, 14609-14620
6. Ma, Q., Akhter, Y., Wilmanns, M., and Ehebauer, M. T. (2014) Active site conformational changes upon reaction intermediate biotinyl-5′ -AMP binding in biotin protein ligase from *Mycobacterium tuberculosis*. *Protein Sci* **23**, 932-939
7. Bagautdinov, B., Kuroishi, C., Sugahara, M., and Kunishima, N. (2005) Crystal Structures of Biotin Protein Ligase from *Pyrococcus horikoshii* OT3 and its Complexes: Structural Basis of Biotin Activation. *J Mol Biol* **353**, 322-333
8. Tron, C. M., McNae, I. W., Nutley, M., Clarke, D. J., Cooper, A., Walkinshaw, M. D., Baxter, R. L., and Campopiano, D. J. (2009) Structural and Functional Studies of the Biotin Protein Ligase from *Aquifex aeolicus* Reveal a Critical Role for a Conserved Residue in Target Specificity *J Mol Biol* **387**, 129-146
9. Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D., and Matthews, B. W. (2006) Co-repressor Induced Order and Biotin Repressor Dimerization: A Case for Divergent Followed by Convergent Evolution. *J Mol Biol* **357**, 509-523
10. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterisation of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci* **22**, 762-773

CHAPTER 6:

DISCUSSION

CHAPTER OUTLINE

This chapter will review the main findings of this thesis regarding the structure of class III BPLs and their potential as novel antifungal targets. The results will be discussed in relation to the current literature, and future directions for this project will also be explored.

6.1 Summary of key findings

Eukaryotic BPLs form a unique structural class (termed class III BPLs) containing a large N-terminal extension in addition to the catalytic domain and C-terminal cap that are common to all BPLs. Structural investigation of class III enzymes is crucial for the targeting of fungal BPLs for the development of anti-infective therapeutics against human and agricultural pathogens; a possibility due to the necessity of BPL for fungal viability (1,2). Human BPL is structurally similar to fungal enzymes, also being a class III enzyme (3). Therefore, a structural understanding of class III BPLs will aid in selective targeting of pathogenic BPLs from all classes over the human homologue, and will also provide insights into different disease mechanisms of multiple carboxylase deficiency (MCD) (reviewed in Chapters 1 and 5).

The BPLs from the prototypical yeast *S. cerevisiae* and the two agricultural fungal pathogens *B. cinerea* and *Z. tritici* demonstrated high levels of sequence similarity together with similar overall structures and tertiary folds. However, subtle structural differences were identified between the enzymes, indicated by their unique stabilities and differing Michaelis constants (K_M) for substrates. These BPLs also demonstrated an ability to prioritise biotinylation of particular biotin domain substrates, shown by the variable Michaelis constants for the biotin domains from different fungal and bacterial species. Interestingly, the three fungal BPLs had different patterns of substrate preference, providing further evidence towards the selective 'substrate validation' mechanism proposed for class III BPLs. Selective and potent *in vitro* inhibition of these three enzymes was demonstrated using small molecule inhibitors that mimic the reaction intermediate. In particular, BPL199 inhibited activity of *Zt*BPL with a nanomolar inhibition constant whilst *Bc*BPL, the most sequence identical and similarly behaved enzyme in all structural, stability and kinetic assays employed, was not inhibited by BPL199 (> 56-fold selectivity) (Table 1). Therefore, despite the homology and overall structural similarity between fungal class III BPLs, there exists the possibility to selectively target individual enzymes to produce narrow spectrum antifungals. The human enzyme has previously been shown to be unaffected by BPL199, allowing selective targeting of pathogenic BPLs over the human isoform and preventing cytotoxic effects to the host (Table 1). Importantly, the potent inhibition of the different class III BPLs by several inhibitors confirmed the potential of BPL as a novel antifungal drug target.

In order to aid further antifungal development against class III BPLs, structural information about this class of enzymes is required. The structure of the BPL from *S. cerevisiae* was

	Human BPL	*Sc*BPL		*Bc*BPL		*Zt*BPL		Cytotoxicity	
	K_i (μM)	K_i (μM)	Fold Selectivity	K_i (μM)	Fold Selectivity	K_i (μM)	Fold Selectivity	HepG2 (μg/mL)	HEK293 (μg/mL)
Biotin acetylene	3.5	2.3	1.5	3	1.2	2.8	1.3	> 64	ND
Biotinol-5′-AMP	0.21	0.186	1.1	0.659	ND	0.063	3.3	> 200	ND
BPL199	> 10	0.02	> 500	> 10	ND	0.178	> 56.2	> 250	> 250
BPL068	> 90	> 10	ND	> 10	ND	> 10	ND	> 64	ND
BPL178	> 16	> 10	ND	> 10	ND	> 10	ND	> 40	ND
BPL223	> 10	> 10	ND	3.9	> 2.6	4	> 2.5	> 80	ND

Table 1. The efficacy and selectivity of biotin and reaction intermediate mimic inhibitors against the class III fungal BPLs from *S. cerevisiae*, *B. cinerea* and *Z. tritici*. These are compared to the *in vitro* inhibition constants of the human BPL (previously reported in the literature) to allow calculations of the selectivity of the different inhibitors for each fungal BPL. Furthermore, the inhibitors have no cytotoxic affects on human liver (HepG2) and kidney (HEK293) cell lines. Data presented are from chapter 4 and the literature (biotin acetylene (4), biotinol-5′-AMP (5), BPL199 (Appendix 2), BPL068 (6), BPL178 (7) and BPL 223 (8)). ND = not determined.

further investigated. HDX and homology modelling revealed a domain in the N-terminal extension that is proposed to contain the structural fold of a glutamine amidotransferase. This is separated from the catalytic domain, which is hypothesised to possess a homologous fold to the catalytic domain of other BPLs, by a partially folded and buried linker region of 160 residues. No homologous structure was identified for this linker, hence, it was not modelled. Structural changes that occur concurrently with ligand binding were revealed. Whilst apo-*Sc*BPL has a more open, dynamic conformation, the binding of ligands rigidifies and compacts the overall enzyme, such that its stability increases. HDX revealed the specific ligand-induced conformational changes were localised both around the active site and on a surface patch on the structured domain of the N-terminal extension. The role of these structural rearrangements and their contribution to allowing class III BPLs to carry out biotinylation remains to be determined.

6.2 Towards an experimentally determined atomic resolution class III BPL structure

Previous attempts to determine the atomic resolution structure of a class III BPL by crystallography or the structure of the N-terminal domain by NMR were not successful (discussed chapter 2). Hence, in this thesis, alternate techniques were investigated to provide structural information. However, the ability to derive a completely experimentally determined atomic resolution structure, which does not rely on predictive modelling, is crucial. The models produced in this thesis were supported with experimental MS techniques, but the ultimate validation will only come from empirical structure determination. New structural biology techniques, advances and improvements in current methods, and the novel structural knowledge presented in this thesis will all contribute towards solving a class III BPL structure.

X-ray crystallography has long been known as the gold-standard technique for structure determination of large proteins. Full length *Sc*BPL has been recalcitrant to crystallisation in the past. However, many different approaches for the crystallisation of large and/or ‘difficult’ proteins have been developed. Some of these approaches that are predicted to be most beneficial for crystallising a class III BPL are discussed below, but this is in no way exhaustive. MS structural techniques together with sequence alignments and homology modelling (chapters 4 & 5) revealed *Sc*BPL is dynamic and contains multiple predicted surface loops. Neither of these features are conducive to successful protein crystallisation (9,10). Whilst the addition of the ligands biotin and ATP were shown to stabilise and reduce dynamic movement of *Sc*BPL (chapter 5), the addition of these has not improved crystallisation success. Continued utilisation of these ligands, together with the following new

approaches, may help further stabilise and structurally order the protein to enable crystal formation. One such technique involves optimising the protein sequence to remove or shorten surface loops and dynamic regions, and prioritise favourable surface residues that are more hydrophilic or contain shorter, less flexible side chains to improve crystal formation. Orthologue screening, which involves attempting to crystallise protein homologs from a variety of species, can achieve this as some sequences may have fewer or shorter loops, be less dynamic, or have favourable surface residues that promote crystal formation (11). The BPLs produced in chapter 4 are currently in crystallography trials, but no reproducible crystals of diffraction quality have been obtained thus far. Other class III BPLs, including those from thermophilic yeast strains could also be screened for their ability to crystallise. Thermophilic proteins are generally more ordered and structured, characteristics that are favourable for crystallography, as this provides increased protein stability at the high temperatures where thermophiles are found (12-14). Another approach for optimising protein sequences for crystallography involves incubating the protein with trace amounts of proteases. Here limited proteolysis removes disordered and/or loop regions, leaving the folded structured domains. This approach has allowed the crystallisation of proteins that were previously intractable (15-17). Protein constructs can also be logically engineered to remove undesirable protein features that may be inhibiting crystal formation. For example, surface loops or disordered regions can be excised and replaced, and unfavourable surface residues (hydrophobic or large mobile side chains) can be mutated to increase protein solubility and provide order to the protein's surface (9,18-25). Protein sequences can also be fused to crystallisation tags in a technique called 'carrier-driven' or 'chaperone-assisted' crystallography. These tags are easily crystallisable proteins that can promote initial crystallisation of the fused partner. This technique has assisted in the successful crystallisation of proteins that have previously not formed crystals (22,26).

Another common approach for large, difficult to crystallise proteins is to truncate the protein into its separate domains (9,27). Whilst the production of an isolated N-terminal domain from the human BPL was previously unsuccessful (28), the homology modelling presented in chapter 5 offers new predictions about the domain boundaries within *Sc*BPL for this approach to be revisited. Residue 240 no longer appears to be a primary domain boundary for the *Sc*BPL N-terminal domain, as was originally inferred from the limited proteolysis data (29), as HDX suggests the following linker sequence is still folded and/or buried. Therefore, a series of protein constructs (potentially fused to solubility tags) with the boundary sequentially extended into the unmodelled linker could be screened for the production of folded, stable

domains for crystallography trials or, alternatively, NMR spectroscopy. The deuterium incorporation (chapter 5, supplementary figure 3) reveals the first surface exposed residues in the linker are at residue 260, with the next major surface exposed region beginning at residue 310. Therefore, constructs comprising residues 1-260 and 1-310 would be obvious starting points for investigation. Likewise, a construct for the expression of the isolated catalytic domain and C-terminal cap could allow successful structural resolution of this part of the enzyme – a strategy not previously attempted. Bacterial BPLs that contain only these domains (ie class I) and have been successfully crystallised, namely *M. tuberculosis, P. horikoshii* and *A. aeolicus* (30-36). A structure of the catalytic domain would confirm the homology of the catalytic domain and C-terminal cap with other BPLs from different species and structural classes. It would also reveal the structure of the ligand-binding site, which is advantageous for antifungal drug design. Furthermore, such a structure would also allow investigation of the biotinylation reaction mechanism and whether the ordered ligand-binding and disordered-to-ordered loop transitions, which are common amongst other species BPLs, are conserved as predicted by HDX. The boundary for this domain is more easily predicted using sequence alignments with previously crystallised BPLs that allow the identification of the corresponding initiating residue of the domain in *Sc*BPL (amino acid 363 through to 690, see chapter 4) or using the domain boundaries predicted from the homology model (381-690) (chapter 5). Developing methods for the individual production of the N-terminal domain and the C-terminal catalytic domain (together with the C-terminal cap) would allow further biophysical experiments to investigate the activities of each domain, and the interactions between each domain and the biotin domain substrate, as well as to use the catalytic domain to study ligand binding, especially for potential antifungal ligands. This strategy is likely to be difficult and the isolation of these proteins as well folded domains may not be possible as the linker region may be a crucial element for the folding and solubility of each domain. Despite this, with the new structural information presented in this thesis, it is worth pursuing. Optimisation of constructs that define the N-terminal and C-terminal domain boundaries, and the choice and localisation of purification, solubility and/or crystallisation tags, will still be required via the trial and error expression of a range of different recombinant protein constructs (18,37). Whilst some servers allow the prediction of the expression of soluble protein and the crystallography success of different constructs (such as XtalPred (38), CRYSTALP2 (39), SECRET (40), CrysPred (41), PXS (42) and SSPF Crystallisation Propensity Predictors (43,44)), ultimately empirical optimisation of these approaches will be required.

Alternative structural techniques that may facilitate the determination of a class III BPL structure continue to emerge and improve. Cryo-electron microscopy (cryo-EM) has lately been cemented as a prominent method for structure determination (45), with the resolution and molecular size of protein structures that can be solved continually improving. Currently, the best resolution achieved is 1.8 Å for a 334 kDa protein (46), and the smallest protein structure solved was 64 kDa (solved to a resolution of 3.2 Å) (47). Theoretically, the molecular size of a class III BPL (76-80 kDa) would enable its structure to be solved by cryo-EM, providing the sample preparation which includes the cryogenic freezing of protein onto grids, which is often the most challenging step of cryo-EM, can be optimised (48). Many aspects of cryo-EM have dramatically improved, and are predicted to continue improving, including the microscopes for imaging, and software for data processing and structure refinement (48,49). Therefore, cryo-EM should be investigated as a technique to derive an atomic resolution structure of a class III BPL.

Another emerging structural technique called micro-ED, which is a hybrid between crystallography and cryo-EM, is looking promising for structure determination of particularly troublesome proteins. Here, electron diffraction patterns produced from small crystals of only a few protein layers (< 500 nM in thickness) are solved to give atomic resolution structures (50). As only small crystals, termed microcrystals, are required, these can often be found directly from crystallisation condition screens, including in microtray wells that contain precipitated protein (50,51). Therefore, optimisation to produce large, regular shaped crystals, necessary for X-ray diffraction, is not required (52). Protein precipitation is commonly identified in the crystallisation screens of fungal class III BPLs, which may provide suitable microcrystals for micro-ED. Likewise, small crystals have been produced for these BPLs, and although they do not diffract X-rays, electrons interact more strongly with matter and can give high quality diffraction data from much smaller crystals (50). Micro-ED has been used to solve protein structures with resolutions between 1 and 3 Å (51), and there is no molecular size restraint (cf. cryo-EM), due to the use of electron diffraction rather than the direct imaging used for cryo-EM. Data analysis for micro-ED is carried out using software tools developed for solving structures from X-ray diffraction patterns and, therefore, this aspect of the technique benefits immensely from the years of development completed with X-ray crystallography (53). As this technique begins to become more commonplace, it may be utilised for the determination of a class III BPL structure. All three of these structural techniques can be employed to determine a 3D atomic resolution protein structure *ab initio* (49,52). However, the homology models presented in this thesis could also be utilised to assist the process of

solving the atomic resolution structure of a BPL by being used as molecular replacement models or being fit to the cryo-EM data. Therefore, the emerging structural techniques of cryo-EM and micro-ED should be considered for the definitive determination of a class III BPL structure.

6.3 Further understanding the role of the N-terminal domain in protein biotinylation

Structural information can help us understand the role of the N-terminal extension and its involvement in protein biotinylation. Previous literature revealed this domain is crucial for BPL activity by facilitating enduring reactions between BPL and the biotin domain substrate (3,29,54,55). However, many of these studies were based on point or truncation mutants of *Sc*BPL or human BPL, the correct folding of which was often not confirmed. A future goal of this project involves confirming the role of the N-terminal extension in enhancing BPL activity by interacting with the biotin domain substrate. Capturing a structure of a class III BPL-biotin domain complex and studying the interactions between enzyme and protein substrate will provide further mechanistic understanding of the function of the N-terminal extension, including whether it assists in the hypothesised 'substrate verification' activity of class III BPLs whereby they select only appropriate substrates for biotinylation.

Multiple attempts to study the complex between *Sc*BPL and *S. cerevisiae* biotin domain (from pyruvate carboxylase isoform 1) using a variety of techniques, including native MS and 2D heteronuclear single quantum coherence (HSQC) NMR, were trialled as part of the experiments for this thesis (data not shown). In each case, the complex was not detected even in the presence of the ligands biotin and MgATP, or the non-hydrolysable reaction intermediate mimic biotinol-5′-AMP. Biotinol-5′-AMP is a potent inhibitor of *Sc*BPL (K_i: 186 nM, chapter 4). Utilising this non-hydrolysable reaction intermediate mimic was hypothesised to promote a more enduring protein: protein complex with the biotin domain substrate, as biotin transfer would be unable to occur. However, this was not observed and the time scale of the interaction was therefore inferred to be too fast to capture the complex by these techniques. SPR has revealed the association and dissociation rates of the human BPL for its biotin domain (from human pyruvate carboxylase) to be 1.9 x $10^3 M^{-1} s^{-1}$ and 2 x $10^{-3} s^{-1}$ respectively, suggesting why the complex could not be captured (3). Covalent crosslinkers that chemically link protein residues localised within defined distances, dictated by the crosslinker length, have also been utilised as part of this study (data not shown) to study the BPL-biotin domain interaction. These crosslinking reactions occur on a very fast time scale, and allow the identification, via MS, of protein regions located in close spatial proximities both

within and between protein molecules (56,57). Hence, protein regions involved in the protein-protein interactions between BPL and the biotin domain can be determined. This method has been utilised to map other protein-protein interactions in the literature (58-60). Another potential approach for studying this complex involves mutating BPL residues involved in catalysis to prevent the biotinylation reaction proceeding. Modification of two specific residues, R48 and K111A, allowed the structure of the *P. horikoshiii* BPL in complex with its biotin domain (from the *P. horikoshii* methylmalonyl-CoA decarboxylase gamma chain) to be solved by X-ray crystallography (32). These residues are involved in binding the triphosphate moiety of ATP, and mutating these residues results in the disordering of the phosphate moiety in the crystal structure such that they are not visible. Consequently, these mutations disrupt the binding of ATP and prevent the formation of the reaction intermediate, catalytically crippling the enzyme (32). Hence, mutating the analogous residues in *Sc*BPL (R425 and K529) may promote the interaction with the biotin domain substrate to allow the study of this interaction via MS, HSQC NMR and X-ray crystallography studies.

Capturing the BPL-biotin domain complex and structurally observing the interaction would provide multiple insights to class III enzymes. In addition to revealing details of class III BPL structure, the structure of the complex would reveal the role the N-terminal domain plays in facilitating biotinylation, particularly identifying residues that interact with the biotin domain or catalytic region of the BPL, or residues that may aid in the catalytic reaction. Mutagenesis of these N-terminal residues could be completed to confirm their function, and specifically whether and how they affect BPL activity. Structural information regarding this interaction and the N-terminal domain itself would also allow MCD-causing mutations to be mapped and elucidate how they facilitate disease, providing insights into novel therapeutic strategies for patients with biotin non-responsive BPL mutations. Finally, interactions between different BPLs and biotin domains (i.e. from different species or biotin-dependent enzymes within a species) can be compared to determine the mechanisms that facilitate the selective ‘substrate validation’ activity apparent in class III BPLs, and whether the N-terminal domain facilitates this selectivity.

6.4 Considerations for developing antifungal therapeutics

BPL is essential for the viability of fungi, and this makes it a promising novel antifungal drug target. New antifungal products are required to overcome devastating pathogens that cause disease to humans and in economically valuable agricultural crops, especially where these

pathogens are developing resistance to the few current antifungals available. Previously BPLs have been exploited as targets for novel antibiotics, but never for antifungals. Further validation of BPL as a target and exploration of biotin biology in different fungal pathogens is crucial to assist drug design efforts. It is necessary to confirm BPL is an essential enzyme in particular fungal pathogens, as so far it has only been confirmed in *S. cerevisiae*, *Schizosaccharomyces pombe* and *C. albicans* (1,2,61,62) (and certain bacteria (34,63,64)). Furthermore, conditional knockouts/downs of BPL in plant or mice models of fungal infections would determine whether BPL inhibition promotes the clearance of infection, and also whether BPL is required for virulence and establishing infections. Some fungi produce their own biotin from specific precursors (65-68); therefore, quantitation of the availability of these precursor molecules during infection is required to determine whether the fungi can synthesise enough biotin to outcompete BPL inhibitors – with this being a recently identified resistance mechanism to BPL inhibitors within *S. aureus* (Appendix 2). Some pathogenic fungi also establish biofilms where they can survive in a latent form protected from the host's immune system (69,70). Fungi in biofilms are often harder to treat due to increased expression of efflux pumps, the impaired penetrance of drugs through the extracellular matrix of biofilms, and the presence of metabolically dormant cells that are often resistant to antifungals (69,71). The necessity of biotin for fungi in biofilms, particularly fungi with altered metabolic states, should be elucidated to determine if BPL is an appropriate target for fungi in this lifestyle. Furthering our understanding of the biotin biology in fungi is crucial for determining strategies to target BPL.

The approach our group, and others, have pursued for pharmacologically targeting bacterial BPLs has been to generate chemical mimics of biotin or biotinyl-5′-AMP (4-6,34,36,72-74). These outcompete the natural substrates to bind within the active site and prevent protein biotinylation. A similar approach can be utilised for fungal BPLs, with cross-reactive biotinyl-5′-AMP mimics shown to be potently active *in vitro* (chapter 4). These inhibitors have previously been thoroughly characterised for their inhibition of the human BPL isoform (Table 1) (4-8) (& Appendix 2), allowing the selectivity of each inhibitor for a pathogenic BPL over the human enzyme to be determined. BPL199 demonstrated the greatest selectivity with 56- and 500-fold greater inhibition for *Zt*BPL and *Sc*BPL respectively over the human and *B. cinerea* isoforms (Table 1). In particular, BPL199 has been shown to have low cytotoxicity against human liver (HepG2) and kidney (HEK293) cells lines (Table 1) (4-8) (& Appendix 2). Whilst sequence alignments (chapter 4) revealed large conservation in the active sites of class III BPLs, with only subtle residue differences, chemical selectivity was still evident between the

three fungal enzymes and the human BPL (particularly between *Bc*BPL and *Zt*BPL that are the most sequence similar enzymes). The extent to which selectivity between these enzymes can be exploited will remain to be investigated with the rational, structure guided development of more potent, whole cell active compounds. If BPL inhibitors are employed to control fungal infections in an agricultural setting, their cross-reactivity with plant BPLs and their toxicity to plants must be evaluated. This would require *in vitro* inhibitor testing against recombinantly produced plant BPLs and *in vivo* cytotoxicity testing against plant tissue, analogous to the tests completed to determine human toxicity. Interestingly, the class I BPL from the prototypical plant *Arabidopsis thaliana* is only modestly similar to the BPLs from humans (24% identity, 40% similarity) and *S. cerevisiae* (22% identity, 40% similarity), suggesting selectivity against plants should also be possible (75). Additionally, BPL inhibitors are designed to bind to an ATP-binding pocket. However, there are many ATP utilising enzymes with possibly similar ATP binding pockets (76) present within human cells. Therefore, these inhibitors must be selective such that they don't inhibit off-target ATP-utilising enzymes, which is common with nucleotide analogue inhibitors such as kinase inhibitors (77,78). Incorporating biotin or a biotin analogue as part of the bi-substrate inhibitor introduces specificity for the BPL over other nucleotide binding sites, as other ATP-binding proteins are unlikely to contain an adjacent pocket that can accommodate a biotin-like moiety, preventing off-target binding. Such a tethering approach has previously been utilised to introduce selectivity for specific kinase targets (78). The biotin-binding pocket is very highly conserved between the BPLs from different species (4). Therefore, this part of the compound is unable to offer species selectivity. Instead, selectivity can be engineered into the v-shaped linker that joins the biotin and adenine moieties of bi-substrate inhibitors, or modifications to the adenine base situated in the ATP-binding site. Chemical modifications to the reaction intermediate at these sites have produced selective inhibitors that preferentially target the *S. aureus* BPL against the human isoform (6).

So far BPL inhibitors have focused on mimicking the reaction intermediate to block the BPL active site, but other parts of the enzyme may be targets for chemical intervention. The emerging evidence of an important role of the N-terminal domain in biotinylation suggests this site could be targeted in fungal BPLs. Small molecules, peptides or other bioactive therapeutics could be developed as allosteric inhibitors to prevent catalytic activity or as protein-protein interface inhibitors to prevent the important interactions of the N-terminal domain with the biotin domain (3,55,79). Sequence alignments (chapter 4) revealed the greatest sequence variation amongst different class III BPLs, particularly between the fungal

and human enzymes, is within the N-terminal extension. Therefore, targeting the N-terminal extension for BPL inhibition may provide a greater potential for selectivity over the human enzyme and between different fungal BPLs to produce inhibitors with narrow-spectrum activity. Plant BPLs do not have this large N-terminal extension (75,80). Hence, targeting this site for the development of agricultural fungicides would provide selectivity for pathogens and reduce toxicity to plants.

Regardless of the inhibitor design approach, or the type of the inhibitor developed, several key aspects must be overcome to develop a bioactive antifungal. Firstly, a BPL inhibitor must traverse the cell wall and membrane to enter the cell and engage with BPL, which is primarily cytoplasmic in eukaryotes (81). Many current fungal drug targets are located in the cell wall or membrane, which overcomes the challenge of cell permeability (82). However, as antifungals with novel mechanisms of action are required to overcome the developing resistance, new targets need to be investigated. Since the 'low hanging fruit' of easy to access external targets are being depleted, novel intracellular targets must be identified. The literature provides various examples of such novel, intracellular targets that are involved in metabolic, signal transduction and gene expression pathways. It also describes the successful development of antifungal inhibitors that can access these intracellular targets, and are in various stages of *in vitro*, preclinical, and phase 1 and 2 clinical trials (83-86). Another important consideration for the development of a bioactive antifungal is the safety and toxicity of the drug candidates, especially as current human therapeutics can have dramatic side effects (82,87). Hence, improvements in the toxicity profiles of antifungals are required. Toxicity and selectivity have been discussed above for both therapeutic and agricultural antifungals, however, there will be extra considerations required for antifungals intended for agricultural use. The toxicity of BPL inhibitors to crops, fruits and other plants will require testing, along with assessing the safety of releasing these compounds into the environment when applied to crops. This will include analysing the safety for consumption by animals and humans, and effects on soil and waterway environments.

The emergence of fungi resistant to current therapeutics is a major reason for the need for new antifungals. As a novel target, BPL should circumvent the existing mechanisms of resistance. However, there is the potential for fungi to develop resistance to BPL inhibitors. *S. aureus* treated with sub-minimal inhibitory concentrations of a BPL inhibitor revealed a spontaneous resistance rate below 1×10^{9} (Appendix 2), whilst a BPL inhibitor targeting *M. tuberculosis* revealed a spontaneous resistance rate of less than 1.4×10^{-10} (74). These

spontaneous resistance rates are encouraging, showing resistance does not develop quickly. These spontaneous resistance rates are lower those of *S. aureus* to the common antibiotics ciprofloxacin, rifampin and mupirocin (88), and better or on par with the resistance rates of *M. tuberculosis* to its common treatments isoniazid and rifampicin (89). Furthermore, resistance to BPL inhibitors is proposed to be less likely to occur in fungi as 5 different biotin-dependent enzymes depend on BPL (90) (compared to 1 or 2 in bacteria (91)). Therefore, multiple important metabolic pathways are inhibited and mechanisms to overcome the inhibition of each pathway would be required for specific resistance to occur. While a resistant mutation in BPL itself would overcome the inhibition of all of these pathways in one step, the necessity of BPL in multiple pathways would provide a large selection pressure for cells to maintain an acitve BPL and, therefore, reduce the likelihood of resistance occurring. Generalised intrinsic resistance mechanisms, such as increased efflux or decreased uptake of the antimicrobial, must still be considered. However, the low spontaneous resistance rates suggest even these mechanisms (i.e. intrinsic and BPL mutations) are not easily acquired. Once lead inhibitors have been identified and developed, and their mechanisms of action confirmed, spontaneous resistance experiments can be completed to reveal the onset of resistance in different fungal pathogens. Combining these experiments with complete genomic sequencing can allow the mechanisms of resistance in these fungi to be identified, revealing potential resistance mechanisms that may emerge from the clinical or agricultural use of these inhibitors. Therefore, ways to overcome or prevent this resistance can be addressed, or combination therapies that provide greater selective pressure to further prevent resistance occurring can be investigated.

6.5 References

1. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Rhonda, B., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., Bakkoury, M. E., Foury, F., Friend, S. H., Gentalen, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D. J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J. L., Riles, L., Roberts, C. J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R. K., Véronneau, S., Voet, M., Volckaert, G., Ward, T. R., Wysocki, R., Yen, G. S., Yu, K., Zimmermann, K., and Philippsen, P. (1999) Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science* **285**, 901-906
2. Hoja, U., Wellein, C., Greiner, E., and Schweizer, E. (1998) Pleiotropic phenotype of acetyl-CoA-carboxylase-defective yeast cells: Viability of a BPL1-amber mutation depending on its readthrough by normal tRNA (Gln)(CAG). *Eur J Biochem* **254**, 520-526
3. Mayende, L., Swift, R. D., Bailey, L. M., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., and Polyak, S. W. (2012) A novel molecular mechanism to explain biotin-unresponsive holocarboxylase synthetase deficiency. *J Mol Med* **90**, 81-88
4. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Zvarec, O., Bell, J. M., Turnidge, J. D., Wallace, J. C., Booker, G. W., Wilce, M. C. J., Abell, A. D., and Polyak, S. W. (2012) Biotin Analogues with Antibacterial Activity Are Potent Inhibitors of Biotin Protein Ligase. *ACS Med Chem Lett* **3**, 509-514
5. Tieu, W., Polyak, S. W., Paparella, A. S., Yap, M. Y., Soares da Costa, T. P., Ng, B., Wang, G., Lumb, R., Bell, J. M., Turnidge, J. D., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2015) Improved Synthesis of Biotinol-5′-AMP: Implications for Antibacterial Discovery. *ACS Med Chem Lett* **6**, 216-220
6. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Pendini, N. R., Polyak, S. W., Pedersen, D. S., Morona, R., Turnidge, J. D., Wallace, J. C., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2012) Selective inhibition of Biotin Protein Ligase from *Staphylococcus aureus*. *J Biol Chem* **287**, 17823-17832
7. Feng, J., Paparella, A. S., Tieu, W., Heim, D., Clark, S., Hayes, A., Booker, G. W., Polyak, S. W., and Abell, A. D. (2016) New Series of BPL Inhibitors To Probe the Ribose-Binding Pocket of *Staphylococcus aureus* Biotin Proteing Ligase *ACS Med Chem Lett* **7**, 1068-1072
8. Paparella, A. S., Lee, K. J., Hayes, A. J., Feng, J., Feng, Z., Cini, D., Deshmukh, S., Booker, G. W., Wilce, M. C. J., Polyak, S. W., and Abell, A. D. (2018) Halogenation of Biotin Protein Ligase Inhibitors Improves Whole Cell Activity against *Staphylococcus aureus*. *ACS Infect Dis* **4**, 175-184
9. Derewenda, Z. S. (2004) The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **34**, 354-363
10. Grimm, C., Klebe, G., Ficner, R., and Reuter, K. (2000) Screening orthologs as an important variable in crystallization: preliminary X-ray diffraction studies of the tRNA-modifying enzyme S-adenosyl-methionine:tRNA ribosyl transferase/isomerase. *Acta Crystallogr D Biol Crystallogr* **56**, 484-488
11. Page, R. (2008) Strategies for improving crystallization success rates. *Methods Mol Biol* **426**, 345-362
12. Jenney Jr, F. E., and Adams, M. W. W. (2008) The impact of extremophiles on structural genomics (and vice versa). *Extremophiles* **12**, 39-50

13. Edwards, T. E., Liao, R., Phan, I., Myler, P. J., and Grundner, C. (2012) Mycobacterium thermoresistibile as a source of thermostable orthologs of *Mycobacterium tuberculosis* proteins. *Protein Sci* **21**, 1093-1096
14. Cava, F., Hidalgo, A., and Berenguer, J. (2009) *Thermus thermophilus* as biological model. *Extremophiles* **13**, 213-231
15. Dong, A., Xu, X., Edwards, A. M., Midwest Centre for Structural Genomics, and Structural Genomics Consortium. (2007) *In situ* proteolysis for protein crystallization and structure determination. *Nat Methods* **4**, 1019-1021
16. Wernimont, A., and Edwards, A. (2009) *In Situ* Proteolysis to Generate Crystals for Structure Determination: An Update. *PLoS One* **4**
17. Gheyi, T., Rodgers, L., Romero, R., Sauder, J. M., and Burley, S. K. (2010) Mass Spectrometry Guided In Situ Proteolysis to Obtain Crystals for X-Ray Structure Determination. *J Am Soc Mass Spectrom* **21**, 1795-1801
18. Müller, I. (2017) Guidelines for the successful generation of protein-ligand complex crystals. *Acta Crystallogr D Struct Biol* **73**, 79-92
19. Tsurumura, T., Qiu, H., Yoshida, T., Tsumori, Y., and Tsuge, H. (2014) Crystallization and preliminary X-ray diffraction studies of a surface mutant of the middle domain of PB2 from human influenza A (H1N1) virus. *Acta Crystallogr F Struct Biol Commun* **70**, 72-75
20. Lawson, A. J., Walker, E. A., White, S. A., Dafforn, T. R., Stewart, P. M., and Ride, J. P. (2009) Mutations of key hydrophobic surface residues of 11β-hydroxysteroid dehydrogenase type 1 increase solubility and monodispersity in a bacterial expression system. *Protein Sci* **18**, 1552-1563
21. Roos, G., Brosens, E., Wahni, K., Desmyter, A., Spinelli, S., Wyns, L., Messens, J., and Loris, R. (2006) Combining site-specific mutagenesis and seeding as a strategy to crystallize 'difficult' proteins: the case of *Staphylococcus aureus* thioredoxin. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 1255-1258
22. Moon, A. F., Mueller, G. A., Zhong, X., and Pedersen, L. C. (2010) A synergistic approach to protein crystallization: Combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci* **19**, 901-913
23. Loll, P. J., Peining, X., Schmidt, J. T., and Melideo, S. L. (2014) Enhancing ubiquitin crystallization through surface-entropy reduction. *Acta Crystallogr F Struct Biol Commun* **70**, 1434-1442
24. Cooper, D. R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M., and Derewenda, Z. (2007) Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr D Biol Crystallogr* **63**, 636-645
25. Longenecker, K. L., Garrad, S. M., Sheffield, P. J., and Derewenda, Z. S. (2001) Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of PhoGDI. *Acta Crystallogr D Biol Crystallogr* **57**, 679-688
26. Wild, R., and Hothorn, M. (2017) The macro domain as fusion tag for carrier-driven crystallisation. *Protein Sci* **26**, 365-374
27. Xu, Y., Guo, J., Wang, L., Jiang, R., Jin, X., Liu, J., Fan, S., Quan, C.-S., and Ha, N.-C. (2016) The Crystal Structure of the YknZ Extracellular Domain of ABC Transporter YknWXYZ from *Bacillus amyloliquefaciens*. *PLoS One* **11**
28. Mayende, L. (2012) *Functional characterisation of the N-terminal region of Holocarboxylase synthetase.* Doctor of Philosophy, University of Adelaide
29. Polyak, S. W., Chapman-Smith, A., Brautigan, P. J., and Wallace, J. C. (1999) Biotin Protein Ligase from *Saccharomyces cerevisiae*. The N-terminal domain is required for complete activity. *J Biol Chem* **274**, 32847-32854
30. Tron, C. M., McNae, I. W., Nutley, M., Clarke, D. J., Cooper, A., Walkinshaw, M. D., Baxter, R. L., and Campopiano, D. J. (2009) Structural and Functional Studies of the

Biotin Protein Ligase from *Aquifex aeolicus* Reveal a Critical Role for a Conserved Residue in Target Specificity *J Mol Biol* **387**, 129-146
31. Bagautdinov, B., Kuroishi, C., Sugahara, M., and Kunishima, N. (2005) Crystal Structures of Biotin Protein Ligase from *Pyrococcus horikoshii* OT3 and its Complexes: Structural Basis of Biotin Activation. *J Mol Biol* **353**, 322-333
32. Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2008) Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate. *J Biol Chem* **283**, 14739-14750
33. Ma, Q., Akhter, Y., Wilmanns, M., and Ehebauer, M. T. (2014) Active site conformational changes upon reaction intermediate biotinyl-5′ -AMP binding in biotin protein ligase from *Mycobacterium tuberculosis*. *Protein Sci* **23**, 932-939
34. Duckworth, B. P., Geders, T. W., Tiwari, D., Boshoff, H. I., Sibbald, P. A., Barry III, C. E., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2011) Bisubstrate Adenylation Inhibitors of Biotin Protein Ligase from *Mycobacterium tuberculosis*. *Chem Biol* **18**, 1432-1441
35. Gupta, V., Gupta, R. K., Khare, G., Salunke, D. M., Surolia, A., and Tyagi, A. K. (2010) Structural Ordering of Disordered Ligand-Binding Loops of Biotin Protein Ligase into Active Conformations as a Consequence of Dehydration. *PLoS ONE* **5**
36. Bockman, M. R., Kalinda, A. S., Petrelli, R., De la Mora-Rey, T., Tiwari, D., Liu, F., Dawadi, S., Nandakumar, M., Rhee, K. Y., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2015) Targeting *Mycobacterium tuberculosis* Biotin Protein Ligase (MtBPL) with Nucleoside-Based Bisubstrate Adenylation Inhibitors. *J Med Chem* **58**, 7349-7369
37. Malawski, G. A., Hillig, R. C., Monteclaro, F., Eberspaecher, U., Schmitz, A. A. P., Crusius, K., Huber, M., Egner, U., Donner, P., and Müller-Tiemann, B. (2006) Identifying protein construct variants with increased crystallization propensity - A case study. *Protein Sci* **15**, 2718-2728
38. Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A., and Godzik, A. (2007) XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403-3405
39. Kurgan, L., Razib, A. A., Aghakhani, S., Dick, S., Mizianty, M., and Jahandideh, S. (2009) CRYSTALP2: sequence-based protein crystallization propensity prediction. *BMC Struct Biol* **9**
40. Smialowski, P., Schmidt, T., Cox, J., Kirschner, A., and Frishman, D. (2006) Will my protein crystallize? A sequence-based predictor. *Proteins* **62**, 343-355
41. Kantardjieff, K. A., and Rupp, B. (2004) Protein isoelectric point as a predictor for increased crystallization screening efficiency *Bioinformatics* **20**, 2162-2168
42. Price II, W. N., Chen, Y., Handelman, S. K., Neely, H., Manor, P., Karlin, R., Nair, R., Liu, J., Baran, M., Everett, J., Tong, S. N., Forouhar, F., Swaminathan, S. S., Acton, T., Xiao, R., Luft, J. R., Lauricella, A., DeTitta, G. T., Rost, B., Montelione, G. T., and Hunt, J. F. (2009) Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat Biotechnol* **27**, 51-57
43. Overton, I. M., Padovani, G., Girolami, M. A., and Barton, G. J. (2008) ParCrys: a Parzen window density estimation approach to protein crystallization propensity prediction. *Bioinformatics* **24**, 901-907
44. Overton, I. M., and Barton, G. J. (2006) A normalised scale for structural genomics target ranking: the OB-Score. *FEBS Lett* **580**, 4005-4009
45. Earl, L. A., Falconieri, V., Milne, J. L. S., and Subramaniam, S. (2017) Cryo-EM: beyond the microscope. *Curr Opin Struct Biol* **46**, 71-78
46. Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R., Boxer, M. B., Earl, L. A., Milne, J. L. S., and Subramaniam, S. (2016) Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **165**, 1698-1707

47. Khoshouei, M., Radjainia, M., Baumeister, W., and Danev, R. (2017) Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. *Nat Commun* **8**
48. Passmore, L. A., and Russo, C. J. (2016) Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol* **579**, 51-86
49. Bai, X.-c., McMullan, G., and Scheres, S. H. W. (2015) How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci* **40**, 49-57
50. Nannenga, B. L., and Gonen, T. (2018) MicroED: a versatile cryoEM method for structure determination. *Emerg Top Life Sci* **2**, 1-8
51. Rodriguez, J. A., Eisenberg, D. S., and Gonen, T. (2017) Taking the measure of MicroED. *Curr Opin Struct Biol* **46**, 79-86
52. Sawaya, M. R., Rodriguez, J., Cascio, D., Collazo, M. J., Shi, D., Reyes, F. E., Hattne, J., Gonen, T., and Eisenberg, D. S. (2016) Ab initio structure determination from prion nanocrystals at atomic resolution by MicroED. *PNAS* **113**, 11232-11236
53. Nannenga, B. L., Bu, G., and Shi, D. (2018) The Evolution and Advantages of MicroED. *Front Mol Biosci* **5**
54. Campeau, E., and Gravel, R. A. (2001) Expression in *Escherichia coli* of N- and C-terminally Deleted Human Holocarboxylase Synthetase: Influence of the N-terminus on biotinylation and identification of a minimum functional protein. *J Biol Chem* **276**, 12310-12316
55. Hassan, Y. I., Moriyama, H., Olsen, L. J., Bi, X., and Zempleni, J. (2009) N- and C-terminal domains in human holocarboxylase synthetase participate in substrate recognition. *Mol Genet Metab* **96**, 183-188
56. Liu, F., and Heck, A. J. R. (2015) Interrogating the architecture of protein assemblies and protein interaction networks by cross-linking mass spectrometry. *Curr Opin Struct Biol* **35**, 100-108
57. Sinz, A. (2018) Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein-Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angew Chem Int Ed Engl* **57**, 6390-6396
58. Smith, D.-L., Götze, M., Bartolec, T. K., Hart-Smith, G., and Wilkins, M. R. (2018) Characterization of the Interaction between Arginine Methyltransferase Hmt1 and Its Substrate Npl3: Use of Multiple Cross-Linkers, Mass Spectrometric Approaches, and Software Platforms. *Anal Chem* **90**, 9101-9108
59. Zybailov, B., Gokulan, K., Wiese, J., Ramanagoudr-Bhojappa, R., Byrd, A. K., Glazko, G., Jaiswal, M., Mackintosh, S., Varughese, K. I., and Raney, K. D. (2015) Analysis of Protein-protein Interaction Interface between Yeast Mitochondrial Proteins Rim1 and Pif1 Using Chemical Cross-linking Mass Spectrometry. *J Proteomics Bioinform* **8**, 243-252
60. Kühn-Hölsken, E., Lenz, C., Dickmanns, A., Hsiao, H.-H., Richter, F. M., Kastner, B., Ficner, R., and Urlaub, H. (2010) Mapping the binding site of snurportin 1 on native U1 snRNP by cross-linking and mass spectrometry. *Nucleic Acids Res* **38**, 5581-5593
61. Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H.-S., Duhig, T., Nam, M., Palmer, G., Han, S., Jeffery, L., Baek, S.-T., Lee, H., Shim, Y. S., Lee, M., kIM, L., Heo, K.-S., Noh, E. J., Lee, A.-R., Jang, Y.-J., Chung, K.-S., Choi, S.-J., Park, J.-Y., Park, Y., Kim, H. M., Park, S.-K., Park, H.-J., Kang, E.-J., Kim, H. B., Kang, H.-S., Park, H.-M., Kim, K., Song, K., Song, K. B., Nurse, P., and Hoe, K.-L. (2010) Analysis of a genome-wide set of gene deletions in the fissuion yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* **28**, 617-623
62. Shtifman Segal, E., Gritsenko, V., Levitan, A., Yadav, B., Dror, N., Steenwyk, J. L., Silberberg, Y., Mielich, K., Rokas, A., Gow, N. A. R., Kunze, R., Sharan, R., and Berman, J. (2018) Gene Essentiality Analyzed by *In Vivo* Transposon Mutagenesis and Machine Learning in a Stable Haploid Isolate of *Candida albicans*. *mBio* **9**

63. Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40
64. Gerdes, S. Y., Scholle, M. D., Campbell, J. W., Balázsi, g., Ravasz, E., Daugherty, M. D., Somera, A. L., Kyrpides, N. C., Anderson, I., Gelfand, M. S., Bhattacharya, A., Kapatral, V., D'Souza, M., Baev, M. V., Grechkin, Y., Mseeh, F., Fonstein, M. Y., Overbeek, R., Barabási, A.-L., Oltvai, Z. N., and Osterman, A. L. (2003) Experimental Determination and System Level Analysis of Essential Genes in *Escherichia coli* MG1655. *J Bacteriol* **185**, 5673-5684
65. Phalip, V., Kuhn, I., Lemoine, Y., and Jeltsch, J.-M. (1999) Characterization of the biotin biosynthesis pathway in *Saccharomyces cerevisiae* and evidence for a cluster containing *BIO5,* a novel gene involved in vitamer uptake *Gene* **232**, 43-51
66. Zhang, S., Sanyal, I., Bulboaca, G. H., Rich, A., and Flint, D. H. (1994) The Gene for Biotin Synthase from *Saccharomyces cerevisiase*: Cloning, Sequencing, and Complementation of *Escherichia coli* Strains Lacking Biotin Synthase. *Arch Biochem Biophys* **309**, 29-35
67. Hall, C., and Dietrich, F. S. (2007) The Reacquisition of Biotin Prototrophy in *Saccharomyces cerevisiae* Involved Horizontal Gene Transfer, Gene Duplication and Gene Clustering. *Genetics* **177**, 2293-2307
68. Ahmad Hussin, N., Pathirana, R. U., Hasim, S., Tati, S., Scheib-Owens, J. A., and Nickerson, K. W. (2016) Biotin Auxotrophy and Biotin Enhanced Germ Tube Formation in *Candida albicans*. *Microorganisms* **4**
69. Gulati, M., and Nobile, C. J. (2016) *Candida albicans* biofilms: development, regulation, and molecular mechanisms. *Microbes Infect* **18**, 310-321
70. Kernien, J. F., Snarr, B. D., Sheppard, D. C., and Nett, J. E. (2018) The Interface between Fungal Biofilms and Innate Immunity. *Front Immunol* **8**
71. Lagree, K., and Mitchell, A. P. (2017) Fungal Biofilms: Inside Out. *Microbiol Spectr* **5**
72. Tieu, W., Jarrad, A. M., Paparella, A. S., Keeling, K. A., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., Polyak, S. W., and Abell, A. D. (2014) Heterocyclic acyl-phosphate bioisostere-based inhibitors of *Staphylococcus aureus* biotin protein ligase. *Bioorg Med Chem Lett* **24**, 4689-4693
73. Shi, C., Tiwari, D., Wilson, D. J., Seiler, C. L., Schnappinger, D., and Aldrich, C. C. (2013) Bisubstrate Inhibitors of Biotin Protien Ligase in *Mycobacterium tuberculosis* Resistant to Cyclonucleoside Formation. *American Chemical Society Medicinal Chemistry Letters* **4**, 1213-1217
74. Bockman, M. R., Engelhart, C. A., Dawadi, S., Larson, P., Tiwari, D., Ferguson, D. M., Schnappinger, D., and Aldrich, C. C. (2018) Avoiding Antibiotic Inactivation in *Mycobacterium tuberculosis* by Rv3406 through Strategic Nucleoside Modification. *ACS Infect Dis* **4**, 1102-1113
75. Tissot, G., Douche, R., and Alban, C. (1997) Evidence for multiple forms of biotin holocarboxylase synthetase in pea (*Pisum sativum*) and in *Arabidopsis thaliana*: subcellular fractionation studies and isolation of a cDNA clone. *Biochem J* **323**, 179-188
76. Denessiouk, K. A., Lehtonen, J. V., and Johnson, M. S. (1998) Enzyme-mononucleotide interactions: Three different folds share common structural elements. *Protein Sci* **7**, 1768-1771
77. Wu, P., Clausen, M. H., and Nielsen, T. E. (2015) Allosteric small-molecule kinase inhibitors. *Pharmacol Ther* **156**, 59-68
78. Gower, C. M., Chang, M. E. K., and Maly, D. J. (2014) Bivalent Inhibitors of Protein Kinases. *Crit Rev Biochem Mol Biol* **49**, 102-115

79. Lee, C.-K., Cheong, C., and Jeon, Y. H. (2010) The N-terminal domain of human holocarboxylase synthetase facilitates biotinylation via direct interaction with the substrate protein. *FEBS Lett* **584**, 675-680
80. Tissot, G., Pepin, R., Job, D., Douche, R., and Alban, C. (1998) Purification and properties of the chloroplastic form of biotin holocarboxylase synthetase from *Arabidopsis thaliana* overexpressed in *Escherichia coli*. *Eur J Biochem* **258**, 586-596
81. Bailey, L. M., Wallace, J. C., and Polyak, S. W. (2010) Holocarboxylase synthetase: Correlation of protein localisation with biological function. *Arch Biochem Biophys* **496**, 45-52
82. Campoy, S., and Adrio, J. L. (2017) Antifungals. *Biochem Pharmacol* **133**, 86-96
83. McCarthy, M. W., Kontoyiannis, D. P., Cornely, O. A., Perfect, J. R., and Walsh, T. J. (2017) Novel Agents and Drug Targets to Meet the Challenges of Resistant Fungi. *J Infect Dis* **216**, S474-483
84. Wiederhold, N. P. (2018) The antifungal arsenal: alternative drugs and future targets. *Int J Antimicrob Agents* **51**, 333-339
85. Li, X., Hou, Y., Yue, L., Shuyuan, L., Du, J., and Sun, S. (2015) Potential targets for antifungal drug discovery based on growth and virulence in *Candida albicans*. *Antimicrob Agents Chemother* **59**, 5885-5891
86. Robbins, N., Wright, G. D., and Cowen, L. E. (2016) Antifungal Drugs: The Current Armamentarium and Development of New Agents. *Microbiol Spectr* **4**
87. Enoch, D. A., Ludlam, H. A., and Brown, N. M. (2006) Invasive fungal infections: a reveiw of epidemiology and management options. *J Med Microbiol* **55**, 809-818
88. Schmitz, F.-J., Fluit, A. C., Hafner, D., Beeck, A., Perdikouli, M., Boos, M., Scheuring, S., Verhoef, J., Köhrer, K., and Von Eiff, C. (2000) Development of Resistance to Ciprofloxacin, Rifampin, and Mupirocin in Methicillin-Susceptible and -Resistant *Staphylococcus aureus* Isolates. *Antimicrob Agents Chemother* **44**, 3229-3231
89. McGrath, M., Gey van Pittius, N. C., van Helden, P. D., Warren, R. M., and Warner, D. F. (2014) Mutation rate and the emergence of drug resistance in *Mycobacterium tuberculosis*. *J Antimicrob Chemother* **69**, 292-302
90. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C. J., Wallace, J. C., and Polyak, S. W. (2008) Biotin protein ligase from *Candida albicans*: Expression, purification and development of a novel assay. *Arch Biochem Biophys* **479**, 163-169
91. Paparella, A. S., Soares da Costa, T. P., Yap, M. Y., Tieu, W., Wilce, M. C. J., Booker, G. W., Abell, A., and Polyak, S. W. (2014) Structure Guided Design of Biotin Protein Ligase Inhibitors for Antibiotic Discovery. *Curr Top Med Chem* **14**, 4-20

APPENDIX 1

CHAPTER OUTLINE

Appendix 1 contains a literature review discussing the potential for targeting biotin protein ligase and biotin biosynthesis enzymes for anti-tuberculosis (anti-TB) agents.

STATEMENT OF AUTHORSHIP

Statement of Authorship

Title of Paper	Biotin Biology as a Target for New Anti-Tuberculosis Drugs
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Published in Jiangsu J Prev Med, May 2016, Vol 27, No. 3, 257-261. doi: 10.13668/j.issn.1006-9070.2016.03.01

Principal Author

Name of Principal Author (Candidate)	Louise Sternicki
Contribution to the Paper	Assisted in preparing the manuscript and figures for publication.
Overall percentage (%)	10%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 6/3/2019.

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Andrew Thompson
Contribution to the Paper	Researched and prepared the manuscript, produced figures for publication.
Signature	Date 8/3/19.

Name of Co-Author	Kate Wegener
Contribution to the Paper	Assisted with preparing the manuscript.
Signature	Date 6/3/19

Name of Co-Author	Wei Lu		
Contribution to the Paper	Assisted with preparing the manuscript. SUPERVISOR.		
Signature		Date	23/5/2019

Name of Co-Author	Li-mei Zhu		
Contribution to the Paper	Assisted with preparing the manuscript. AS SUPERVISOR.		
Signature		Date	23/5/2019

Name of Co-Author	Grant Booker		
Contribution to the Paper	Assisted with preparing the manuscript.		
Signature		Date	6/3/2019.

Name of Co-Author	Steven Polyak		
Contribution to the Paper	Assisted with preparing the manuscript.		
Signature		Date	6/3/2019

Name of Co-Author	Yan Li		
Contribution to the Paper	Assisted with preparing the manuscript. AS SUPERVISOR.		
Signature		Date	23/5/2019

APPENDIX 1

·结核病专题论著·

生物素——抗结核药物新研究方向

THOMPSON Andrew P[1], STERNICKI Louise M[1], WEGENER Kate L[1],
陆伟[2], 竺丽梅[2], BOOKER Grant W[1], POLYAK Steven W[1], 李燕[2]
1. 澳大利亚阿德莱德大学生命科学院, 阿德莱德 5005; 2. 江苏省疾病预防控制中心, 南京 210009

摘要: 耐药菌的出现对全球公共卫生构成了巨大的挑战, 结核分枝杆菌(TB)是结核病的病原菌, 是临床上最重要的病原体之一。新抗生素的研发用以治疗耐药性结核病已成为临床及公共卫生管理亟待解决的重要问题。生物素(维生素 B7)在结核杆菌的代谢途径如脂肪酸合成和三羧酸循环中发挥重要作用, 通过生物素蛋白连接酶连接活性辅助因子而被激活。生物素蛋白连接酶及其相关生物合成通路是微生物合成生物素过程中必不可或缺的关键分子, 因而有可能成为新的药物靶点。遗传研究表明, 结核杆菌需要由菌体进行生物素的全程合成, 因此, 生物素合成酶在在结核杆菌代谢过程中起着至关重要的作用, 引起了抗结核药物研究者的高度重视。本文论述了结核杆菌生物素合成通路, 并重点分析了生物素合成过程中可能成为药物靶点的关键酶, 旨在为结核病的治疗提供新思路, 具有重要的临床意义。

关键词: 结核分枝杆菌; 耐药菌; 生物素; 药物靶点
中图分类号: R52　　**文献标识码**: A　　**文章编号**: 1006－9070(2016)03－0257－05

Biotin Biology as a Target for New Anti-Tuberculosis Drugs

THOMPSON Andrew P*, STERNICKI Louise M, WEGENER Kate L, LU Wei,
ZHU Li-mei, BOOKER Grant W, POLYAK Steven W, LI Yan
*University of Adelaide, School of Biological Sciences, Adelaide 5005, Australia

Abstract: The rise of drug resistant bacteria poses a great global healthcare challenge. One of the most clinically important pathogens is *Mycobacterium tuberculosis*, the agent responsible for tuberculosis (TB). New classes of antibiotics with novel mechanisms of action are required to combat drug resistant TB. The biotin biocycle presents multiple targets for antimicrobial drug discovery, namely the biotin-dependent enzymes biotin protein ligase and the biotin biosynthesis pathway. Biotin (vitamin B7) is a cofactor for enzymes that play key roles in important metabolic pathways in TB, such as fatty acid synthesis and replenishing the citric acid cycle. The biotin dependent-enzymes are activated via the attachment of the cofactor through the activity of biotin protein ligase. Genetic studies suggest that *M. tuberculosis* fulfils its biotin requirements solely through *de novo* synthesis, making the biotin synthetic enzymes attractive for drug discovery. Here we review the biotin biocycle and highlight studies investigating the key enzymes as possible drug targets. This knowledge underpins research efforts to exploit these in the development of new treatments for TB.

Key words: Mycobacterium tuberculosis; Drug resistant bacteria; Biotin; Drug targets

1　Tuberculosis and Drug Resistance

Tuberculosis (TB) is an infectious disease primarily caused by the pathogenic bacterium *Mycobacterium tuberculosis*. TB is a major global health care issue, with an estimated one third of the world's population infected by the latent organism[1]. The risk of becoming ill with TB after bacterial infection is only 10%, but because the number of people infected with the pathogen is so large, TB is ranked alongside HIV/AIDS as a leading cause of deaths worldwide. In 2014, 9.6 million people contracted TB, resulting in 1.5 million deaths[1]. *M. tuberculosis* colonizes the lungs of infected patients, and is transmitted via droplets in the air when infected people cough, sneeze or spit. It is a highly resilient bacterium capable of inhabiting cells in the lung[2] or macrophages where it can survive in a dormant state for extended periods of time[3] and subsequently reactivate to cause TB[2]. Its unique cell wall structure, consisting primarily of mycolic acids, is thought to contribute to its pathogenicity[4-5].

Treatment for TB is a six-month drug regimen-requiring front line antibiotics, such as isoniazid and rifampicin[1]. The extended length of treatment makes patient compliance a major issue, especially if appropriate guidance and support are not available. Poor patient adherence and inappropriate use of the drugs are cited as major factors for the development

DOI: 10.13668/j.issn.1006－9070.2016.03.01
基金项目: National Health and Medical Research Council of Australia (APP1068885); National Natural Science Foundation of China (No. 81301448); Key Medical Talent Foundation of Jiangsu Provincial Center for Disease Control and Prevention (No. JKRC20110029); Jiangsu health international exchange program sponsorship (JSH－2015－002).
作者简介: Thompson Andrew P(1991－), 男, 澳大利亚人, 博士在读, 主要从事以生物素合成酶为靶点的新型药物筛选工作。
通讯作者: Polyak Steven W, E-mail: steven.polyak@adelaide.edu.au; 李燕, E-mail: liyan.nju@163.com

of multidrug resistant TB (MDR-TB). MDR-TB is caused by *M. tuberculosis* strains that are unresponsive to isoniazid and rifampicin. This form of TB requires second line treatments that often have greater associated toxicity. *M. tuberculosis* strains that are resistant to a larger selection of anti-TB drugs, including second line treatments, have recently been documented; these have been termed extensive drug resistant TB (XDR-TB). Approximately 480 000 people contracted MDR-TB in 2014, with an estimated 9.7% of those cases being XDR-TB.

2 Biotin biology and therapeutic intervention

The rise of drug resistant *M. tuberculosis* highlights a need for safer and more effective treatment options. Biotin biosynthesis and utilisation pathways provide an array of potential drug targets for new antimicrobial therapies that could fulfil this need. Biotin is an essential vitamin that is utilised by all living organisms. It is a well-studied enzyme cofactor that is required for certain carboxylation, decarboxylation and transcarboxylation reactions. *M. tuberculosis* possesses two classes of biotin-dependent enzymes that play key metabolic roles in fatty acid metabolism and energy production, namely Acyl-CoA carboxylases (ACCs) and pyruvate carboxylases (PC). These enzymes are activated when ligated to biotin by the enzyme biotin protein ligase (BPL). Here we discuss biotin-dependent enzymes, BPL and biosynthetic enzymesas possible drug targets for the treatment of TB.

2.1 Biotin-dependent enzymes ACCs are a class of enzyme that catalyse the carboxylation of short chain acyl-CoA substrates[6]. The products of these reactions are used in the synthesis of fatty acids and polyketides, including the pathogenically important mycolic acids[4-5]. Mycolic acids are unique to mycolata species, and are a major component of the complex cell wall of *M. tuberculosis*. Multiple anti-TB drugs currently on the market work by inhibiting lipid synthesis, such as isoniazid, ethionamide and thiocarlide, demonstrating the efficacy of such an approach[7]. Isoniazid and ethionamide target an enoyl-acyl carrier protein reductase known as InhA, preventing the synthesis of fatty acids including mycolic acid[8]. Similarly, thiocarlide inhibits the production of oleic acid[9]. The ACCs as targets for new antibiotics have recently been reviewed[10].

PC catalyses the conversion of pyruvate to oxaloacetate, an intermediate of the citric acid cycle[7], and therefore plays a pivotal role in the energy metabolism of cells. *M. tuberculosis* has three oxaloacetate producing enzymes, and the efficacy of inhibiting PC for the treatment of TB has not yet been confirmed[7].

2.2 Biotin protein ligase (BPL) All organisms contain between one and five biotin-dependent enzymes, which each require BPL for biotin ligation[11-12]. Biotin is attached to a specific lysine residue in the active site of the biotin-dependent enzymes. Biotinylation proceeds via a two-step reaction, requiring both biotin and ATP. These reactants first bind to their adjacent binding pockets, yielding the reaction intermediate biotinyl-5′-AMP. The BPL: biotinyl-5′-AMP complex then interacts with the biotin-dependent enzyme such that the target lysine side chain enters the BPL active site. The positively charged ε-amino group of the lysine is then selectively biotinylated[12-13]. Biotin-dependent enzymes require biotinylation for their activity, and without it they are unable to fulfill their important metabolic roles. Therefore, small molecule inhibitors that target BPL represent possible treatments for TB.

Validation of BPL as a drug target has been provided by both genetic and pharmacological studies. Knockout studies showed that deletion of the BPL gene in fast growing model species *Mycobacterium smegmatis*, prevented colony growth, demonstrating that BPL is essential for viability *in vitro*[14]. Two non-hydrolysable analogues of the biotinyl-5′-AMP reaction intermediate have also been shown to pharmacologically inhibit bacterial growth[14-15]. Importantly, these compounds are also potent antibacterials against various drug-resistant *M. tuberculosis* strains including MDR-TB and XDR-TB[14]. Together these studies identify *M. tuberculosis* BPL as an attractive target for antibiotic development.

Certain bacterial BPLs, including the *M. tuberculosis* enzyme, possess an ordered binding mechanism. This knowledge impacts upon the design of inhibitors that target BPL. Here biotin binds to the enzyme first, causing a disordered to ordered conformational change within the enzyme. Restructuring of a biotin-binding loop, positions the side chain of a key tryptophan residue such that it is conducive for ATP binding (figure 1). Binding of both biotin and ATP is required for the synthesis of biotinyl-5′-AMP. Restructuring of a second disordered adenylate-binding loop further stabilizes the interaction with the reaction intermediate[12,16-19]. As the ATP pocket is unstructured without biotin, it cannot be targeted on its own. Therefore, efforts to design BPL inhibitors that target the active site must contain a biotinyl functional group, and this serves as a starting point for the design of new compounds[15-16,20-21].

2.3 Biotin biosynthesis Organisms acquire biotin via a variety of different mechanisms. In humans and other mammals, biotin is obtained through exogenous

APPENDIX 1

sources such as diet or from biotin-producing bacteria in the gut[22-23]. Certain organisms have the metabolic pathways required for the synthesis of biotin, including some plants, fungi, bacteria and other microorganisms[24]. Bacteria can also scavenge extracellular biotin through biotin ATP-dependent transporters. These include BioY, YigM and MadN, which have been characterised in various bacteria[25]. Genome annotation studies and nucleotide sequence analysis with homology algorithms have failed to identify BioY, YigM or MadN homologues in *M. tuberculosis*[7,26-27]. These findings suggest that *M. tuberculosis* relies on *de novo* synthesis as its sole source of biotin, highlighting this pathway as a potential drug target.

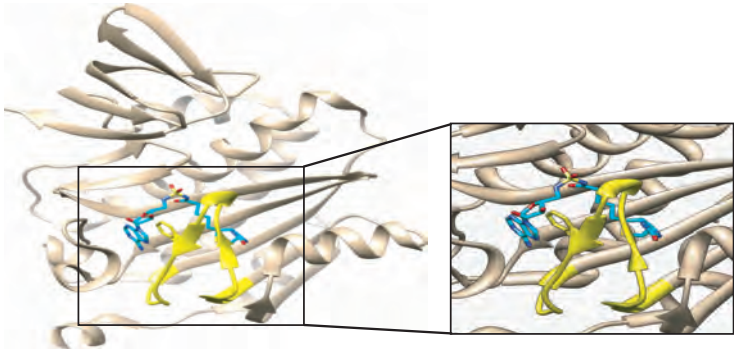


Figure 1 Structure of *M. tuberculosis* BPL with a bisubstrate inhibitor (cyan) that mimics the reaction intermediate (PDB: 3RUX)[14]. This demonstrates adjacent binding pockets for biotin (right) and ATP (left). The biotin-binding loop (yellow) becomes structured after biotin binding, allowing the ATP pocket to form and repositioning a Tryptophan residue (side chain shown) that is necessary for ATP binding. This mechanism results in the ordered binding of biotin prior to ATP. For details in picture printed in colors, please see back cover.

The biotin biosynthesis pathway is responsible for the enzymatic conversion of a malonyl CoA precursor to biotin (figure 2). Briefly, the biotin biosynthetic pathway begins with *O*-methyltransferase (encoded by *bioC*) converting malonyl CoA to malonyl CoA methyl ester, followed by elongation of the alkyl chain using the fatty acid synthesis pathway. Pimeloyl-ACP methyl ester is exited from the fatty acid synthesis pathway and carboxylesterase (encoded by *bioH*) converts it to pimeloyl-ACP. The next four reactions are conserved between all biotin producing organisms. 7-keto-8-aminopelargonic acid (KAPA) is produced by KAPA synthase (encoded by *bioF*) through the use of L-alanine. KAPA is subsequently converted to 7-8-diaminopelargonic acid (DAPA) by DAPA synthase (encoded by *bioA*), utilising S-adenyl methionine (SAM) as an amino donor. Dethiobiotin synthetase (DTBS, encoded by *bioD*) then utilises CO_2 to close the ureido ring of DAPA in an ATP dependent manner to form dethiobiotin[28]. Finally, a sulphur and two electrons from flavodoxin, SAM and nicotinamide adenine dinucleotide phosphate (NADPH) are utilised by biotin synthase to catalyse the closure of the thiophane ring, resulting in the formation of d-biotin. For a more detailed explanation of this metabolic pathway we encourage readers to see our recent review[7].

Several studies provide strong genetic and biochemical evidence for the inhibition of biotin biosynthesis as a strategy for developing anti-TB agents. Genetic knockout studies using strains of *M. tuberculosis* deficient in the *bioA* and *bioF* genes are unable to colonisemice in models of infection[29-30]. A conditional *bioA* knockout reveals that this gene is also essential for persistence as a chronic infection[29]. Similarly, genetic studies in *M. smegmatison* the biotin biosynthesis genes *bioA*, *bioF* and *bioB* confirm that the biotin biosynthesis pathway is essential for *in vitro* growth and *in vivo* pathology[31-33]. Finally, knockout of a reductase required for the synthesis of the biotin precursor pimeloyl-thioester, reduced growth of *Mycobacterium marinum* on blood agar, unless supplemented with high levels of biotin[34]. These studies indicate biotin biosynthesis is a promising drug target for TB infection.

Biotin biosynthesis offers many attractive features as an anti-TB drug target. Foremost of these is the lack of a biotin biosynthetic pathway in mammals. The absence of human homologues means that it is less likely that the drug will have unintended side effects, as commonly seen with current anti-TB drugs. In addition, *M. tuberculosis* is unable to scavenge biotin from the surrounding environment and is solely dependant on *de novo* synthesis. Finally, the biotin synthesis pathway appears to be essential for *M. tuberculosis* survival in the dormancy state. Agents capable of targeting the latent bacterial stage are highly desired as most antibiotics are ineffective against this stage of its lifecycle. Treatments that are active against this stage would allow infections to be resolved before becoming symptomatic.

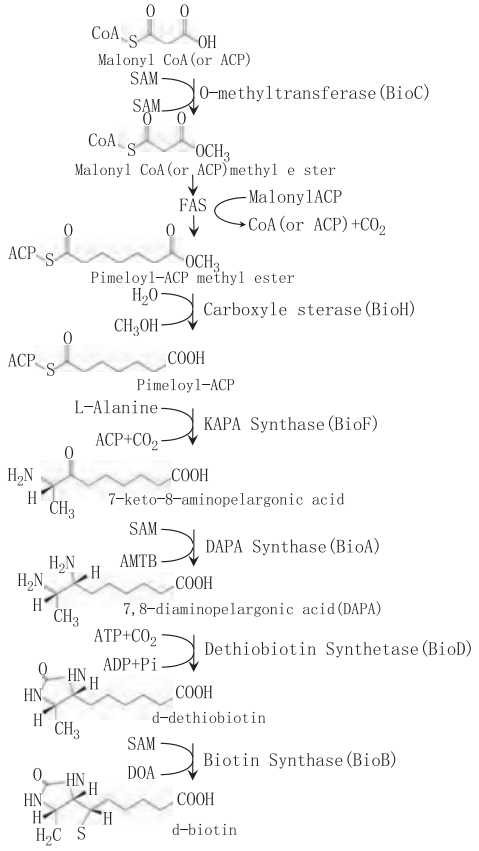


Figure **2** Schematic of the substrates and products of the biotin biosynthesis pathway, with the catalysing enzyme and its encoding gene indicated. Adapted from Salaemae et al[7]

3 Antibiotic Discovery Approaches

The rise of antibiotic resistance constitutes a major health concern, especially considering the lack of antibiotic drug development in recent years. Front line drugs with efficacy against drug resistant strains and that are safer than current treatments are desperately required. Treatments that also target dormant *M. tuberculosis* are also needed. Towards this end, we must discover new agents with novel mechanisms of action, that will not subject to existing resistance mechanisms. The biotin biocycle provides several promising new and highly attractive drug targets for TB research. However, translation of our knowledge about this pathway into novel drugs is still challenging. Antibiotic discovery should be viewed as a highly specialised area research where the rules that govern development of drugs for other diseases areas do not necessarily apply. A recent review from Astra Zeneca detailed their extensive antibiotic drug discovery programs from the last decade, from which no successful drugs were produced[25]. This review highlighted many issues with current antibiotic drug discovery strategies, culminating in the conclusion that these approaches are often not well suited to antibiotic discovery. High rates of attrition arose predominantly at the step of converting a molecule with bioactivity against a purified protein into an antibacterial with whole cell activity. This was due, in part, to poor understanding of how compounds interact with microbes, such as the mechanisms of cell wall penetration and efflux[34]. Also, medicinal chemistry approaches to chemical optimisationcan be problematic, including a tendency to engineer compounds with higher *in vitro* potency that coincides with higher lipophilicity. Antibiotics have been shown to occupy a different chemical spacefrom other drugs and do not abide by the rules and metrics that define traditional 'drug-like' molecules[35]. Consequently, compound libraries that have been curated for other disease areas are often not ideal for antibacterial research. The challenge now is to assemble new libraries with the appropriate chemical diversity for antibiotic discovery[36]. Fragment based libraries, for example, can be investigated as these cover large regions of chemical space. The small molecule compounds obtained from these screens provide starting points for medicinal chemistry to develop more drug-like compounds through chemical extension or concatenation. This structure-guided approach is yet to be fully applied to antibiotic discovery. Natural product libraries also have strong potential, with the biological world already providing many important chemical scaffolds with potential as antibacterial drugs, such as the recent discovery of secondary metabolite texiobactin[37]. Drug screening projects using novel and diverse compound libraries, and that contain bioactive compounds that target biotin biosynthesis and its utilisation, promise exciting new avenues for TB drug discovery. Researchers in China and around the world continue to search niche environments for new bioactive compounds, for example in the South China Sea or from plants used in traditional Chinese medicine. Towards this end, Chinese researchers in partnership with international collaborators can have an impact on the world stage. It is through multidisciplinary research that we have the best opportunity to finally combat TB.

Acknowledgements: We are grateful to the Wallace and Carthew families for their financial support of this work.

Reference

[1] World Health Organisation. Global tuberculosis report[R]. Geneva: WHO, 2015.

[2] SINGH G, SINGH G, JADEJA D, et al. Lipid hydrolizing enzymes in virulence: Mycobacterium tuberculosis as a model system[J]. Crit Rev Microbiol, 2010, 36(3): 259.

[3] PODINOVSKAIA M, LEE W, CALDWELL S, et al. Infection of macrophages with Mycobacterium tuberculosis induces global modifications to phagosomal function: Mtb infection modifies phagosome function[J]. Cell Microbiol, 2013, 15(6): 843.

[4] TAKAYAMA K, WANG C, BESRA GS. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis [J]. Clin Microbiol Rev, 2005, 18(1): 81.

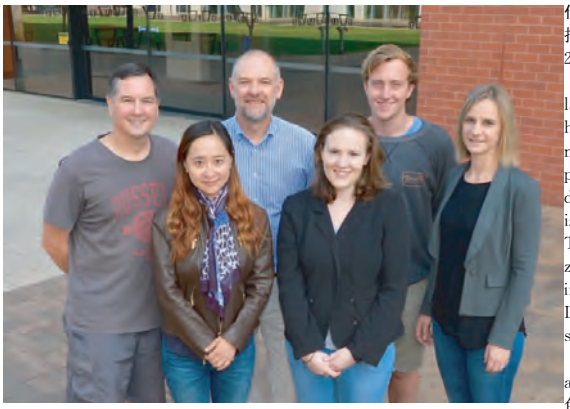
[5] BRENNAN PJ, CRICK DC. The cell-wall core of Mycobacterium tuberculosis in the context of drug discovery[J]. Curr Top Med Chem, 2007, 7(5): 475.

[6] SALAEMAE W, AZHAR A, BOOKER GW, et al. Biotin biosyn-

APPENDIX 1

thesis in Mycobacterium tuberculosis: physiology, biochemistry and molecular intervention[J]. Protein Cell, 2011, 2(9): 691.
[7] SALAEMAE W, BOOKER GW, POLYAK SW. The role of biotin in bacterial physiology and virulence: a novel antibiotic target for Mycobacterium tuberculosis[J]. Microbiolspec, 2016, 4: 2.
[8] MARRAKCHI H, LANÉELLE G, QUÉMARD A. InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II[J]. Microbiology, 2000, 146(Pt 2): 289.
[9] PHETSUKSIRI B, JACKSON M, SCHERMAN H, et al. Unique mechanism of action of the thiourea drug isoxyl on Mycobacterium tuberculosis[J]. J Biol Chem, 2003, 278(52): 53123.
[10] POLYAK SW, ABELL AD, WILCE MC, et al. Structure, function and selective inhibition of bacterial acetyl-coa carboxylase [J]. Appl Microbiol Biotechnol, 2012, 93(3): 983.
[11] CHAPMAN-SMITH A, CRONAN JE. Molecular biology of biotin attachment to proteins[J]. J Nutr, 1999, 129(Suppl2): 477.
[12] PAPARELLA AS, SOARES DA COSTA TP, YAP MY, et al. Structure guided design of biotin protein ligase inhibitors for antibiotic discovery[J]. Curr Top Med Chem, 2014, 14(1): 4.
[13] BAGAUTDINOV B, MATSUURA Y, BAGAUTDINOVA S, et al. Protein biotinylation visualized by a complex structure of biotin protein ligase with a substrate[J]. J Biol Chem, 2008, 283 (21): 14739.
[14] DUCKWORTH BP, GEDERS TW, TIWARI D, et al. Bisubstrate adenylation inhibitors of biotin protein ligase from Mycobacterium tuberculosis[J]. Chem Biol, 2011, 18(11): 1432.
[15] TIEU W, POLYAK SW, PAPARELLA AS, et al. Improved synthesis of biotinol-5′-AMP: implications for antibacterial discovery [J]. ACS Med Chem Lett, 2015, 6(2): 216.
[16] SOARES DA COSTA TP, TIEU W, YAP MY, et al. Selective inhibition of biotin protein ligase from Staphylococcus aureus [J]. J Biol Chem, 2012, 287(21): 17823.
[17] MA Q, AKHTER Y, WILMANNS M, et al. Active site conformational changes upon reaction intermediate biotinyl-5′-AMP binding in biotin protein ligase from Mycobacterium tuberculosis [J]. Protein Sci, 2014, 23(7): 932.
[18] WEAVER LH, KWON K, BECKETT D, et al. Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator[J]. Proc Natl Acad Sci U S A, 2001, 98(11): 6045.
[19] WOOD ZA, WEAVER LH, BROWN PH, et al. Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution[J]. J Mol Biol, 2006, 357(2): 509.
[20] SOARES DA COSTA TP, TIEU W, YAP MY, et al. Biotin analogues with antibacterial activity are potent inhibitors of biotin protein ligase[J]. ACS Med Chem Lett, 2012, 3(6): 509.
[21] TIEU W, JARRAD AM, PAPARELLA AS, et al. Heterocyclic acyl-phosphate bioisostere-based inhibitors of Staphylococcus aureus biotin protein ligase[J]. Bioorg Med Chem Lett, 2014, 24 (19): 4689.
[22] POLYAK SW, BAILEY LM, AZHAR A, et al. Biotin (Vitamin H or B7)in: Micronutrients: Sources, properties and health benefits[M]. Nova Science Publishers, New York, 2012: 65.
[23] ZEMPLENI J, WIJERATNE SS, HASSAN YI. Biotin[J]. Biofactors Oxf Engl, 2009, 35(1): 36.
[24] ROJE S. Vitamin B biosynthesis in plants[J]. Phytochemistry, 2007, 68(14): 1904.
[25] AZHAR A, POLYAK SW, BOOKER GW. Mechanisms of biotin transport[J]. Biochem Anal Biochem, 2015, 04(04): 210.
[26] HEBBELN P, RODIONOV DA, ALFANDEGA A, et al. Biotin uptake in prokaryotes by solute transporters with an optional ATP-binding cassette-containing module[J]. Proc Natl Acad Sci U S A, 2007, 104(8): 2909.
[27] RODIONOV DA, MIRONOV AA, GELFAND MS. Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea[J]. Genome Res, 2002, 12(10): 1507.
[28] SALAEMAE W, YAP MY, WEGENER KL, et al. Nucleotide triphosphate promiscuity in Mycobacterium tuberculosis dethiobiotin synthetase[J]. Tuberculosis (Edinb), 2015, 95(3): 259.
[29] WOONG PARK S, KLOTZSCHE M, WILSON DJ, et al. Evaluating the sensitivity of Mycobacterium tuberculosis to biotin deprivation using regulated gene expression[J]. PLoS Pathog, 2011, 7(9): e1002264.
[30] SASSETTI CM, RUBIN EJ. Genetic requirements for mycobacterial survival during infection[J]. Proc Natl Acad Sci U S A, 2003, 100(22): 12989.
[31] KEER J, SMEULDERS MJ, GRAY KM, et al. Mutants of *mycobacterium smegmatis* impaired in stationary-phase survival[J]. Microbiology, 2000, 146(Pt 9): 2209.
[32] SASSETTI CM, BOYD DH, RUBIN EJ. Comprehensive identification of conditionally essential genes in mycobacteria[J]. Proc Natl Acad Sci U S A, 2001, 98(22): 12712.
[33] RENGARAJAN J, BLOOM BR, RUBIN EJ. Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages[J]. Proc Natl Acad Sci U S A, 2005, 102 (23): 8327.
[34] YU J, NIU C, WANG D, et al. MMAR_2770, a new enzyme involved in biotin biosynthesis, is essential for the growth of Mycobacterium marinum in macrophages and zebrafish[J]. Microbes Infect, 2011, 13(1): 33.
[35] TOMMASI R, BROWN DG, WALKUP GK, et al. ESKAPEing the labyrinth of antibacterial discovery[J]. Nat Rev Drug Discov, 2015, 14(8): 529.
[36] COOPER MA. A community-based approach to new antibiotic discovery[J]. Nat Rev Drug Discov, 2015, 14(9): 587.
[37] LING LL, SCHNEIDER T, PEOPLES AJ, et al. A new antibiotic kills pathogens without detectable resistance[J]. Nature, 2015, 517(7535): 455.

Received: 2016-03-02 **Editor**: Peng Hai-yan



作者合影：第一排从左到右为李燕，Sternicki Louise M，Wegener Kate L；第二排从左到右为 Polyak Steven W，Booker Grant W，Thompson Andrew P。于 2016 年 4 月 20 日摄于澳大利亚阿德莱德大学。

Booker Grant W is professor of biochemistry at the University of Adelaide, Australia. Together with Dr Polyak Steven W and Dr Wegener Kate L, he runs an antibacterial drug discovery program with a focus on new treatments for Staphylococcus aureus and Mycobacterium tuberculosis. Their approach is exploit enzymes that require the important micronutrient biotin as drug targets for the discovery of novel classes of antibiotics. The team specializes in structure guided drug design and fragment based drug discovery. Thompson Andrew P is a PhD candidate working on biotin biosynthetic enzymes as a new drug target for TB. Sternicki Louise M's graduate project is investigating biotin protein ligase as a drug target for new anti-infectives. Dr Li Yan was the recipient of a Jiangsu Health International Exchange Program sponsorship to work on an early stage drug discovery at the University of Adelaide.

Booker Grant W 教授，澳大利亚阿德莱德大学生物化学系教授，与 Polyak Steven W 博士和 Wegener Kate L 博士组成科研团队，共同承担治疗金黄色葡萄球菌和结核分枝杆菌的新型抗生素研发项目。研究主要以生物素相关酶类作为药物靶点筛选新型抗生素，生物素是微生物的重要生长因子，正常代谢不可或缺的关键物质。该研究团队擅长药物化学结构改造和小分子新药研发工作。Thompson Andrew P 为在读博士研究生，主要从事以生物素合成酶为靶点的新型药物筛选工作。Sternicki Louise M 为在读博士研究生，主要从事以生物素蛋白连接酶为靶点的新型药物筛选工作。李燕博士为江苏卫生国际交流项目资助的访问学者，主要从事结核病诊断方法和结核病防治相关研究；现在阿德莱德大学从事以生物素蛋白连接酶为靶点的新型抗生素早期研究的相关工作。

APPENDIX 2

CHAPTER OUTLINE

The manuscript presented in this chapter describes the inhibition of the bacterial *S. aureus* BPL by a reaction intermediate mimic. This inhibitor (BPL199) was utilised in chapter 4 to assess the inhibition of fungal BPLs. Here, the dual mechanism of action of BPL199 against *S. aureus* is investigated, along with the potential resistance rates and mechanisms that *S. aureus* may use to overcome inhibition by BPL199. Native MS, previously optimised in chapter 3 for *Sa*BPL, has been employed in this manuscript to elucidate the oligomeric state of the mutant BPL developed by *S. aureus* as a resistance mechanism against the inhibitor. These data, together with other biochemical assessment, were essential in understanding how this mutant BPL facilitates resistance.

STATEMENT OF AUTHORSHIP

Statement of Authorship

Title of Paper	Probing the mechanism of action and resistance of BPL inhibitors in *Staphlococcus aureus* using an antibacterial sulfonyl-based mimic of the reaction intermediate, biotinyl-5'-AMP
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Louise Sternicki
Contribution to the Paper	Prepared samples for and completed native mass spectrometry experiments, analysed MS data and prepared MS figures for publication, prepared manuscript.
Overall percentage (%)	15%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 6/3/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Andrew J Hayes
Contribution to the Paper	Performed antibacterial susceptibility and cytotoxicity assays, completed mechanism of action studies, and advanced resistance experiments. Expressed, purified and characterised *Sa*BPL D200E (together with Jiulia Satiaputra). Assisted with manuscript preparation.
Signature	Date 22/04/19

Name of Co-Author	Jiulia N Satiaputra
Contribution to the Paper	Characterised DNA binding of wildtype *Sa*BPL in response to biotin and inhibitor binding via RT-qPCR, EMSA and LacZ reporter assays. Further characterised DNA binding of *Sa*BPL D200E by EMSA and LacZ reporter assays. Assisted with data analysis and manuscript preparation.
Signature	Date 15/05/19

Name of Co-Author	Ashleigh S Paparella
Contribution to the Paper	Expression and purification of *Sa*BPL proteins, assisted with *in vitro* biotinylation assays.
Signature	Date 28/5/2019

Name of Co-Author	Zikai Feng
Contribution to the Paper	Assisted with spontaneous resistance experiments
Signature	AS SUPERVISOR Date 23/5/2019

Name of Co-Author	Beatrice B Rodriguez
Contribution to the Paper	Synthesised inhibitors.
Signature	AS SUPERVISOR Date 23/5/2019.

Name of Co-Author	Jacko Feng
Contribution to the Paper	Synthesised inhibitors.
Signature	AS SUPERVISOR Date 23/5/2019

Name of Co-Author	William Tiau
Contribution to the Paper	Synthesised inhibitors.
Signature	Date 13/6/2019

Name of Co-Author	Andrew D Abell
Contribution to the Paper	Supervised medicinal chemistry and inhibitor synthesis.
Signature	Date 13/6/2019

Name of Co-Author	Bart Eijelkamp
Contribution to the Paper	Provided assistance with the qRT-PCR assays and their data analysis. Assisted with manuscript preparation.
Signature	Date 6/5/2019.

Name of Co-Author	Keith E Shearwin
Contribution to the Paper	Provided assistance with the *in vivo* reporter assays and their data analysis. Assisted with preparing the manuscript.
Signature	Date 6/3/2019.

Name of Co-Author	Tara L Pukala
Contribution to the Paper	Assisted with mass spectrometry data collection and analysis.
Signature	Date 6/3/2019

Name of Co-Author	Grant W Booker
Contribution to the Paper	Supervised the study, assisted with manuscript preparation.
Signature	Date 6/3/2019

Name of Co-Author	Steven W Polyak
Contribution to the Paper	Conceived and supervised the study, data analysis, prepared the manuscript.
Signature	Date 6/3/2019

MANUSCRIPT

Probing the mechanisms of action and resistance of BPL inhibitors in *Staphylococcus aureus* using an antibacterial sulfonyl based mimic of the reaction intermediate, biotinyl-5′-AMP

Hayes, A.J.[1, 3*], Satiaputra, J.[1, 4*], Sternicki, L.M.[1], Paparella, A.S.[1, 5], Feng, Z.[1, 6], Rodriguez, B.B.[2], Feng, J. [2], Tieu, W.[2, 7], Eijkelkamp, B.A.[1], Shearwin, K.[1], Pukala, T.L.[2], Abell, A. D.[2], Booker, G.W.[1] & Polyak, S.W.[1, 8]

[1] School of Biological Sciences, University of Adelaide, South Australia 5005, Australia
[2] School of Physical Sciences, University of Adelaide, South Australia 5005, Australia
[3] Present address: School of Biomedical Sciences, The Peter Doherty Institute for Infection and Immunity, The University of Melbourne and The Royal Melbourne Hospital, Melbourne, Victoria 3000, Australia.
[4] Present address: Faculty of Health and Medical Sciences, UWA Centre for Medical Research, University of Western Australia, Perth, Western Australia, 6009 Australia
[5] Present address: Albert Einstein College of Medicine, New York City, New York, 10461 United States of America
[6] Present address: School of Natural Sciences and School of Medicine, The University of Tasmania, Hobart, Tasmania, 7001 Australia
[7] Present address: South Australian Health and Medical Research Institute, Adelaide, South Australia, 5000 Australia
[8] Present address: School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, South Australia, 5001 Australia

*These authors contributed equally to this work

Abstract

Biotin protein ligase (BPL) inhibitors are a novel class of antibacterial agent that can target clinically important methicillin-resistant *Staphylococcus aureus.* In certain bacteria such as *S. aureus,* BPL is a bifunctional protein that is responsible for both catalysing protein biotinylation as well as serving as a transcriptional repressor that regulates the genes involved in biotin homeostasis. Whilst BPL inhibitors with *in vitro* antibacterial activity have previously been described, the full effects of these agents on *S. aureus* have never been experimentally explored. Here we investigate the mechanisms of action and resistance for a potent BPL inhibitor with antibacterial activity. The compound was shown to both inhibit the catalytic activity of BPL as well function as a co-repressor *in vivo* to promote transcriptional repression. The spontaneous resistance rate to this antibacterial was low ($< 1 \times 10^{-9}$), and advanced resistance studies led to the evolution of a missense mutation in BPL (D200E) that reduced DNA binding affinity *in vitro,* and transcriptional repression *in vivo*. It is proposed that this resistance mechanism promotes the availability of biotin by derepressing its synthesis and import, such that biotin can outcompete the inhibitor for binding to BPL. This is the first *in vivo* evidence of a BPL inhibitor acting as a co-repressor to promote transcriptional repression and provides new insight into the molecular mechanism of antibacterial activity and the potential for resistance in *S. aureus*.

Introduction

The need for new antimicrobial agents to combat the growing threat of resistance is currently not being met, with only 5 novel classes of antibiotics introduced since 2000 (1,2). Strains that are now resistant to all classes of clinically approved antibiotics have been reported for *Staphylococcus aureus, Enterococci* species and *Mycobacterium tuberculosis* (3). Without new antimicrobial agents, the inability to treat resistant microbial infections is predicted to result in 10 million deaths annually by 2050 (4). Hence, there is a desperate need for new products with novel mechanisms of action that are not subject to existing resistance mechanisms. Precious few antimicrobial agents with a novel mechanism of action are currently in the drug development pipeline (2), with high risk and low financial reward discouraging investment in this area (1,5,6). One promising new drug target is biotin protein ligase (BPL) for which new potent, selective inhibitors are being developed (reviewed (7,8)). Here we examine one such inhibitor that is active against drug resistant *Staphylococcus aureus* and probe mechanisms of action and potential resistance.

In certain bacteria, such as clinically important *S. aureus*, BPL is a bifunctional protein with two important roles; 1) to catalyse the enzymatic biotinylation of biotin-dependent enzymes, and 2) the biotin-dependent repression of certain target genes implicated in biotin homeostasis. This bi-functionality makes BPL the key master regulator of all biotin-mediated metabolic processes in *S. aureus* and a promising new antibiotic target (8). *S. aureus* BPL (*Sa*BPL) catalyses the covalent attachment of a biotin prosthetic group onto two biotin-dependent enzymes, namely acetyl-CoA carboxylase (ACC) and pyruvate carboxylase (PC). Protein biotinylation occurs through two partial reactions. Biotin (**1**) is first ligated to ATP yielding the reaction intermediate biotinyl-5′-AMP (**2**) (**Figure 1a**). BPL then employs the labile biotinyl-5′-AMP to attach the biotinyl moiety onto a single target lysine residue present in the active site of ACC or PC (9) (**Figure 1b**). The appended biotin is necessary for the biotin-dependent enzymes to fulfil their important metabolic roles. ACC catalyses the first committed step in fatty acid synthesis, whilst PC replenishes the TCA cycle with oxaloacetate. ACC and BPL have been identified as essential gene products for *in vitro* growth in genetic knockout studies performed on *S. aureus* (5,6,10,11) and, consequently, have been the targets of drug discovery programs. The antibacterial efficacy of ACC inhibitors has been demonstrated *in vivo* (reviewed (12,13)) and BPL inhibitors have shown efficacy *in vitro* against *S. aureus* and *M. tuberculosis* (14-18). Whilst dispensable for growth in culture, PC has been shown to be an important virulence factor in bacteraemia models in mice and nematodes (19,20). As protein biotinylation is essential for the activity of ACC and PC,

pharmacological inhibition of *S. aureus* BPL (*Sa*BPL) is proposed to impact bacterial viability and pathogenesis due to the disruption of both enzymes (reviewed (8)).

*Sa*BPL also regulates the transcription of three genetic elements, namely the biotin biosynthesis operon (*bioO*), a gene encoding the substrate specific subunit of the biotin transport protein BioY, as well as a putative fatty acid ligase operon *yhfT-yhfS* (21,22). Recent *in vivo* characterisation in *S. aureus* has revealed a hierarchy in the control of these genes by *Sa*BPL. The biotin biosynthesis operon can be potently repressed by 10 nM biotin 15 minutes post-treatment with exogenous biotin (22). Transcripts for *bioY* and *yhfT-yhfS* persist longer after biotin treatment. It is proposed that this mechanism may assist with metabolic adaptation of the bacteria to their environment, as *de novo* biotin biosynthesis is an energetically expensive process that can be bypassed when exogenous biotin is readily available for import. This was supported experimentally by demonstrating that the growth rate of *S. aureus* was enhanced in media containing exogenous biotin even though the bacteria were amenable to its synthesis (22). We propose this ability is important during infection as *S. aureus* occupies a variety of niche microhabitats with varying biotin concentrations (22). These findings also suggest that pharmacological induction of transcriptional repressor activity may lead to biotin starvation through the dual action of reduced *de novo* biotin synthesis and import. Hence, small molecules that serve as co-repressors are beneficial in the design of antibacterials that target *Sa*BPL.

Structural biology has greatly assisted in understanding the molecular basis of biotin-induced transcriptional repression. *Sa*BPL contains an N-terminal helix-turn-helix domain necessary for DNA binding (reviewed (9,23)). When the cellular demand for biotin is low, and biotinyl-5′-AMP is not required for protein biotinylation, the reaction intermediate also serves as a co-repressor by inducing protein homodimerization - a pre-requisite for high affinity DNA-binding to the operator sequences present in the promoter regions of the target operons and genes (24,25). A series of well-characterized conformational changes occur in *Sa*BPL to facilitate this protein dimerization (26). Following the binding of biotin, residues 118–129 (known as the biotin-binding loop) undergo a disordered-to-ordered transition that creates a pocket to accommodate ATP. An adenosine-binding loop formed by residues 224-228, also undergoes structural changes to stabilize the binding of the adenylate moiety (26). Further allosteric conformational changes occur to loops located in the dimerization interface upon ligand binding. These include the ordering of loops contained between amino acids 142 to 150 (extending an alpha-helix by 3 residues) and 198 to 201 (26). Together these structural

rearrangements facilitate the transition from monomer to homodimer. R122 and F123 in the biotin-binding loop stabilise the dimer by interacting directly with the side chain of D200 in the opposing subunit (25). Homo-dimerization appropriately aligns the two N-terminal DNA-binding domains to facilitate their interaction with DNA (26). Protein dimerization has been reported to be induced by chemical analogues of biotin and biotinyl-5′-AMP(14,25,27,28)

Several groups have reported on the development of BPL inhibitors with anti-bacterial activity (29-39). Bi-substrate inhibitors that occupied both the biotin- and ATP-binding pockets and mimic the mode of binding adopted by biotinyl-5′-AMP have been particularly promising. Non-hydrolysable analogues of the reaction intermediate have been generated by replacing the labile phosphoanhyride linker connecting the biotin and AMP moieties with alternative bioisosteres. These new inhibitors exploited intermolecular interactions with the pathogenic BPLs that are devoid with the human enzyme to provide necessary selectivity. Recently, the bisubstrate analogue **3** (**Figure 1c**) was reported with potent anti-*S. aureus* activity (MIC RN4220 = 0.125 µg/mL, ATCC 49775 = 0.25 µg/mL), necessary *in vitro* selectivity over the human equivalent (K_i *Sa*BPL= 2.4 nM, K_i human BPL > 10 µM, > 4000-fold selectivity) (40). This inhibitor contains a non-hydrolysable sulfonyl-based linker to replace the phosphodiester linker present in biotinyl-5′-AMP. The similarity of the chemical structures of **2** and **3** suggested that **3** may possess the same co-repressor activity as biotinyl-5′-AMP. Here we investigate transcriptional responses in *S. aureus* following treatment with **3**. Potential resistance mechanisms were also investigated through advanced resistance studies, and a missense mutation in *Sa*BPL was identified and characterised. Altogether, this study aims to further understand the therapeutic potential of BPL inhibitors, as a step towards developing novel antibiotics with a defined mechanism of action.

Materials and methods

General bacteria culture and molecular biology reagents: All chemicals and reagents were purchased as analytical grade or higher. Unless otherwise stated, all bacterial strains were purchased from the American Tissue Culture Collection. All molecular biology enzymes (restriction enzymes and DNA polymerase) and buffers were acquired from New England Biolabs. Oligonucleotides were purchased from Geneworks Ptd Ltd. Details of strains, plasmids, and oligonucleotides are supplied in the supplementary information.

Quantification of gene expression using QRT-PCR: The qRT-PCR protocols used to measure the expression of *bioD* and *bioY* were completed as outlined previously (22). However, the sub-cultured *S. aureus* NCTC 8325 cultures were treated with either 10 nM biotin or compound **3** at 10 nM or 3.9 µM [i.e. 4 × MIC] once reaching mid-log phase, prior to cell lysis, RNA extraction and qRT-PCR analysis.

Electrophoretic Mobility Shift Assay (EMSA): The EMSA protocol was adapted from methods previously published (22,24,41). For analysis of compound **3**, 100 µM of compound was added to the EMSA buffer. As compound **3** was reconstituted in 100% DMSO, the final DMSO concentration in the binding buffer was 2.5% (v/v). Control reactions were prepared by adding 100 µM biotin and 2.5% (v/v) DMSO to the binding reaction.

Spontaneous resistance rate: A culture containing 10^9 CFU of *S. aureus* NCTC 8325 was plated onto Cation Adjusted Mueller Hinton broth (CAMHB) Agar containing 4 x MIC of compound **3**. Nine technical replicate cultures were plated from the same starting colony to control for priming mutations in the initial colony. Plates were assessed for growth after 24 hours at 37 °C. As the determination of true resistance rates using fluctuation analysis requires some colonies to grow (42-44), a resistance frequency was instead calculated.

Generation of resistant mutants by serial passaging: Mutants were selected by serial passage in CAMHB in 96-well plates as previously described (45). Briefly, a 2-fold dilution series of compound **3** was plated across 7 wells. Initial concentrations ranged from 1 µg/ml to 0.016 µg/ml, with these concentrations increasing (up to 64 µg/ml) as the MIC increased. Approximately 10^4 CFU of *S. aureus* NCTC 8325 at mid log phase were added to each well and allowed to grow at 37 °C for 20 hours. The OD at 620nm was measured and the highest concentration of **3** that permitted growth (OD_{620} > 0.1 used as threshold) was diluted 1000-

fold to inoculate a new concentration series of **3**. This process was repeated for 18 days, and on the final day single isolates were recovered by streaking out the strains from the highest concentrations of **3** that still allowed growth. Genomic DNA was isolated with the Wizard Genomic DNA purification kit (Promega, Wisconsin, USA) according to the manufacturers' guidelines. Bacteria were treated with 10 µg of lysostaphin (Sigma- Aldrich Inc, MO, USA) and 1 mg/ml lysozyme prior to lysis. These selected strains were also propagated for 6 days on CAMHB agar in the absence of antibiotic selection, before re-testing these strains for antibacterial susceptibility to assess the stability of the generated resistance. The *bpl* gene from each of these strains was sequenced to allow the identification of missense mutations potentially providing resistance to **3**.

Protein methods: Expression and purification of non-liganded (ie apo) *Sa*BPL was performed following previously described protocols (24,46,47). Apo-purification was performed by incubating bacterial lysate with streptavidin agarose at 30 °C for 1 hour prior to incubating with purified biotin domain-GST fusion protein from *S. aureus* (48) (*Sa*PC90) at 30 °C for 1 hour, as described previously (24,25). Protein concentration was determined by Bradford assay. Apo-purified protein was confirmed by native nano-electrospray ionisation-mass spectrometry (nESI-MS) and by Western blot analysis as described previously (24,25).

***In vitro* Biotinylation assay and measurement of inhibition constants:** An activity assay utilising time-resolved fluorescence to measure the biotin incorporation by BPL was completed as previously reported (29).

Native nESI-MS: Native MS of apo-*Sa*BPL was completed according to (24). MS parameters were as previously reported to maintain non-covalent interactions and included; capillary voltage, 1.5-1.7 kV; cone voltage, 40-80 V; trap collision energy, 20-50 V; transfer collision energy, 15-20 V; source temperature, 50°C; extraction cone, 2.0-5.0 V; trap gas flow, 5-8 mL/min; backing pressure, 3.95 mbar. Data analysis were performed in MassLynx V4.1 using manual peak finding.

Chromosomal integration and β-galactosidase reporter assay: *Sa*BPL D200E was cloned into the established integration vector (methods provided in supplementary materials) and incorporated chromosomally into the *E. coli* reporter strains prepared previously (22,24), using established methods (49). β-galactosidase assays were performed as previously reported (22,24,49).

Results

Biological evaluation of 3

The antibacterial spectrum, therapeutic window and mechanism of action was first established using a series of microbiological assays. Antimicrobial susceptibility assays performed against a panel of clinical isolates of *S. aureus* (n=23) revealed **3** was active against both methicillin sensitive and resistant strains with MICs ranging from 0.125-0.5 µg/ml (**Table 1**). There was no change to the MIC when the antimicrobial susceptibility assay was performed in growth media supplemented with 10% foetal calf serum but did increase 8-fold with 20% serum suggesting a slightly reduced efficacy in the presence of serum fatty acids (**Supplementary Information Figure S1**). This potent antibacterial activity was restricted to Staphylococci as other bacterial pathogens were insensitive to **3** including Gram-positive *Streptococcus pneumoniae,* and Gram-negative *Escherichia coli* and Enterobacter species. Only *M. tuberculosis* exhibited susceptibility to **3** albeit with weaker sensitivity than *S. aureus* (MIC = 55 µg/ml, 100 µM). Cytotoxicity testing against two human cell lines HepG2 (liver derived) and HEK293 (kidney derived) showed both cell lines were insensitive to **3** at concentrations up to 250 µg/ml, providing a >100 fold therapeutic window for *S. aureus* (**Table 1**, **Supplementary Information Figure S2**).

Two lines of evidence supported the mechanism of action for **3** was through binding to the BPL target. Firstly, the compound showed reduced antibacterial activity against a strain of *S. aureus* engineered to overexpress the BPL target (18). The increased concentration of BPL in the bacteria increased the MIC by 8-fold over a control strain (**Supplementary Information Figure S3a)**. This response was specific for the BPL inhibition with similar shifts observed with another chemical analogue of the reaction intermediate, biotinol-5′-AMP but not for the β-lactam amoxicillin (**Supplementary Information Figure S3b and c**). In further support, increased concentrations of the BPL substrate biotin also antibacterial activity of **3**. Super-physiological concentrations of 1 µM (ie >1000-fold serum concentrations) of exogenous biotin increased the MIC of **3** by 64-fold (**Supplementary Information Figure S4**). Finally, the concentration and time dependent kill kinetic experiments were performed by viable cell counts following 24 hours of treatment with **3** at 1, 2 and 4× MIC. Growth was completely perturbed at all concentrations, however, killing was not observed consistent with **3** functioning as a bacteriostatic antibacterial (**Supplementary Information Figure S5**). The bacteriostatic activity of **3** means it requires either host induced killing or combination therapy to clear infection. Consequently synergy assays with other clinically used antibacterial agents

was performed using a checkerboard assay (50). Of the 8 antibacterial agents tested none were antagonistic with **3**, and synergy was detected with both methicillin and streptomycin as determined by a fractional inhibitory concentration (FIC) of <0.5 (50,51). (**Supplementary Information Figure S6**). Together these biological data help establish **3** as a promising new antibacterial agent.

Inhibitor 3 induces conformational changes required for DNA binding and functions as an active co-repressor *in vivo*

The *Sa*BPL repressor binds to DNA as a homo-dimer, with each monomer binding to one of the two half-site sequences in the target promoter. Hence, the dimer assembles such that the two N-terminal helix-turn-helix motifs are positioned appropriately for DNA binding. To investigate if inhibitor **3** may function as a co-repressor the published co-crystal structure of *Sa*BPL in complex with **3** (PDB ID: 6ORU) (40) was compared with that of the enzyme bound to biotinyl-5′-AMP (**2**) (PDB ID: 3RIR) (26). In both instances *Sa*BPL crystallised as a homodimer in the crystallographic unit cell. Superimposition of the two structures revealed high structural similarity (RMSD = 1.2 Å) (**Figure 2**). Both ligands **2** and **3** bound to *Sa*BPL in the same conformation, adopting the necessary V-shaped geometry required to fill the adjacent biotin and ATP-binding pockets in the active site (**Figure 2**, **Supplementary Information Figure S7**). These structural data also revealed that binding of **3** induced the same conformational changes in *Sa*BPL as noted for the reaction intermediate that are required for homodimerization. This includes the positioning of residues F123, R122 and D200 within the dimerization interface. Importantly, the overlay analysis also revealed that the N-terminal domains were positioned in same orientation suggesting both structures were receptive to DNA binding. This activity was confirmed using *in vitro* electrophoretic mobility shift assays (EMSA). The affinities of the *Sa*BPL-**3** and *Sa*BPL-biotinyl-5′-AMP complexes binding to the *bioO* and *bioY* promoters were compared. Prior to EMSA analysis, apo-*Sa*BPL was incubated with either compound **3**, or biotin and MgATP to allow the formation of biotinyl-5′-AMP**.** The *Sa*BPL-**3** complex exhibited equipotent DNA binding to both the *bioY* and *bioD* promoter sequences as the *Sa*BPL-biotinyl-5′-AMP complex (**Figure 3**). Complete binding of 10 nM *bioO* required 156 nM of *Sa*BPL independent of whether biotinyl-5′-AMP or compound **3** were bound (**Figure 3a-b**). For 10 nM of *bioY*, a higher concentration of 312 nM *Sa*BPL was required to bind 100% of the oligonucleotide for both ligands (**Figure 3c-d**). Together these structural and *in vitro* data demonstrate that **3** induces protein dimerization and DNA-binding in *Sa*BPL.

We next addressed if **3** could indeed function as a co-repressor *in vivo*. *S. aureus* NCTC 8325 was grown in biotin depleted media until log-phase then treated with 10 nM of either biotin or **3**. Cells were also treated with 3.9 µM of **3** which was 4-times the minimal inhibitory concentration. Cells were harvested at 0, 15, 30 and 90 minutes post treatment and RNA extracted for qRT-PCR analysis to analyse the levels of the *bioD* (the first gene in the biotin biosynthesis operon) and *bioY* (biotin transporter gene) transcripts (**Figure 4**). *S. aureus* treated with 10 nM biotin displayed a 111-fold decrease in *bioD* expression at 15 minutes post-treatment and maintained this level of tight repression throughout the 90-minute time course (**Figure 4a**). Addition of 10 nM of **3** only resulted in a 14-fold repression of *bioD* expression 15 minutes post addition ($p \leq 0.0001$, **Figure 4b**). Similarly, this level of repression was maintained throughout the 90-minute time course. Stronger repression of *bioD* was achieved by increasing the dosage of **3** to 3.9 µM. A 52-fold decrease in *bioD* transcripts was measured at 15 minutes, and *bioD* expression was further repressed over the time course with a 32-fold decrease at 30 minutes and 71-fold decrease at 90 minutes ($p \leq 0.0001$, **Figure 4c**). The *bioY* transcript exhibited less dramatic repression than *bioD* when either biotin or **3** were added. With the addition of 10 nM biotin, *bioY* was only repressed 2-fold at 15 minutes ($p < 0.01$), 5–fold at 30 minutes ($p < 0.01$) and 6-fold at 90 minutes ($p < 0.01$) (**Figure 4d**). *S. aureus* treated with 10 nM of **3** exhibited a similar response, with *bioY* expression reduced by 1.5-fold at 15 minutes ($p < 0.01$), 2-fold at 30 minutes ($p < 0.01$) and 6-fold at 90 minutes ($p \leq 0.0001$, **Figure 4e**). These data suggest **3** can function as a co-repressor, promoting transcriptional repression comparable to that of the natural co-repressor biotinyl-5′-AMP.

Generation of *S. aureus* mutants resistant to 3

The ability of *S. aureus* to develop resistance to **3** was investigated using both spontaneous resistance and serial passage approaches. The *S. aureus* NCTC 8325 strain was utilised due to the availability of a sequenced reference genome for comparative sequence analysis. To measure the rate at which spontaneous resistance develops, 10^9 *S. aureus* cells were plated onto media containing 4x the MIC of **3**. No resistant mutants grew after 24 hours leading to a calculated resistance frequency of $< 10^{-9}$ at this inhibitor concentration. Resistant *S. aureus* strains were subsequently generated by the serial passage of seven populations of *S. aureus* NCTC 8325 in sub-inhibitory concentrations of **3** for 18 days. After 18 days of passage, individual colonies were isolated. These were then passaged on non-selective media to verify that the inherited mutations were genetically stable. The *bpl* gene from selected strains was sequenced to determine if there were amino acid substitutions in the enzyme that may provide resistance to compound **3**. Aspartic acid at residue 200 was substituted with glutamic

acid (D200E) in one clone. Analysis of the available structural data revealed this mutation resides in the dimerization interface between the two BPL subunits. We subsequently explored the proposal that *Sa*BPL-D200E may have altered dimerization activity.

_Sa_BPL D200E acts by disruption of dimerization and, hence, transcriptional repression

Site directed mutagenesis was employed to re-engineer the D200E substitution into the protein sequence from *S. aureus* Mu50 for which structural data have been published. Recombinant wildtype and *Sa*BPL-D200E were purified in their non-liganded (apo) state for further *in vitro* and *in vivo* characterisation. Mass spectrometry (MS) confirmed the desired mutation was present in *Sa*BPL-D200E (measured mass: 37904.96 Da; predicted mass: 37904.79 Da). Native nano-electrospray ionisation (nESI) MS and a previously described Streptavidin- blot method both verified that the both proteins were purified in their apo-states with no ligands bound (**Figure 5, Supplementary Information Figure S8**). The catalytic properties of wildtype and D200E *Sa*BPL were first measured using an *in vitro* protein biotinylation assay. The biochemical data revealed that enzyme activity was not significantly affected by the D200E mutation, with the affinity for biotin increased modestly by 2-fold (wild-type *Sa*BPL K_M 1.8 ± 0.3 µM, *Sa*BPL D200E K_M 3.8 ± 0.4 µM ($p < 0.05$)) (**Table 2**). Furthermore, the D200E mutation had little effect upon the inhibitory activity of **3** (K_i wild-type *Sa*BPL 4.8 ± 2.1 nM; K_i SaBPL-D200E = 10.9 ± 3.5 nM; $p = 0.18$) (**Table 2**). Native nESI-MS was then utilised to investigate the oligomeric state of *Sa*BPL D200E. This technique has previously been employed to study co-repressor induced dimerization of *Sa*BPL (24). Whilst the wildtype protein underwent a well characterized monomer-to-dimer transition induced by biotinyl-5′-AMP (**Figure 5**), *Sa*BPL D200E was monomeric in both its apo and holo forms (**Figure 5, Supplementary Information Table 1**). Therefore, this mutation prevented dimerization when the reaction intermediate was bound, even at high protein concentrations of 90 µM. The native nESI-MS provided sufficient mass resolution to confirm the presence of bound biotinyl-5′-AMP, indicating that lack of protein dimerization is not due to the inability to catalyse the co-repressor.

The effect of the impaired homodimerization ability on DNA-binding and transcriptional repression were then assessed using *in vitro* EMSA analysis and a previously established *in vivo* transcriptional reporter system in *E. coli* (22). Both wild-type and *Sa*BPL D200E demonstrated similar weak DNA binding activity in their non-liganded (ie apo) states in an EMSA. A minimum *Sa*BPL concentration of 20 nM was required to detect binding to the *bioO* oligonucleotide, and 39 nM of enzyme required to detect binding to the *bioY* oligonucleotide

(**Figure 6a-d**). Both proteins failed to completely bind all *bioO* and *bioY* probes at the protein and oligonucleotide concentrations (10 nM) tested, as indicated by the detection of unbound DNA oligonucleotides. The EMSA analysis was then repeated in the presence of biotin and MgATP to measure DNA-binding in response to co-repressor. Unlike the apo protein, 50 nM of *Sa*BPL was sufficient to bind all of the oligonucleotide, highlighting the higher affinity for DNA associated with binding of the corepressor. Furthermore, the minimum concentrations of holo-wild-type *Sa*BPL required to bind DNA (3.12 nM) were 6-fold and 12-fold lower than the apo protein for *bioO* and *bioY* respectively, consistent with previous studies that have shown DNA can enhance homo-dimerization (24,25) (**Figure 6e,g**). However, the mutation had a greater impact upon DNA binding. Holo-*Sa*BPL D200E only demonstrated detectable DNA binding at a higher protein concentration of 10 nM for the *bioO* oligonucleotide and 19 nM for the *bioY* oligonucleotide (**Figure 6f,h**). The protein concentrations to detect DNA binding of holo-*Sa*BPL D200E were lower than those required to detect DNA binding of apo-*Sa*BPL D200E. Whilst complete binding of the 10 nM of both *bioO* and *bioY* oligonucleotides were observed with 50 nM of holo-wild-type *Sa*BPL (**Figure 6e,g**), approximately 3-fold higher concentrations of holo-*Sa*BPL D200E were required for the same response with *bioO* (156 nM) (**Figure 6f**). Holo-*Sa*BPL D200E failed to achieve complete binding to the *bioY* oligonucleotide, even at the highest protein concentration tested (156 nM) (**Figure 6g,h**). These data revealed that the D200E mutation not only impaired BPL dimerization, but also reduced the affinity of *Sa*BPL for DNA.

Finally, the biological consequence of the D200E mutation upon repressor activity was determined using an *in vivo* reporter assay. This *in vivo* reporter assay was previously developed in a biotin auxotrophic *E. coli* strain that had the DNA-binding activity of its endogenous BPL removed (49). The assay was adapted such that the transcriptional repressor activities of both wild-type *Sa*BPL and *Sa*BPL D200E could be investigated against the *S. aureus bioO* and *bioY* promoters. The expression of a β-galactosidase reporter gene was measured in the presence of varying biotin concentrations, and a repression constant determined ($K_{R\ biotin}$ i.e. the biotin concentration required for a 50% reduction of β-galactosidase activity). A 3.2 fold increase ($p < 0.05$) in biotin concentration was required for repression of the *bioO* promoter between wild-type *Sa*BPL ($K_{R\ biotin} = 4.3 \pm 1.9$ nM) and the D200E mutant ($K_{R\ biotin} = 13.9 \pm 3.4$ nM) (**Figure 7a**, **Table 3**). A larger difference was observed for the *bioY* promoter, where wild-type BPL ($K_{R\ biotin} = 8.2 \pm 0.7$ nM) repressed at vastly lower biotin concentrations than *Sa*BPL D200E ($K_{R\ biotin} > 500$ nM), which did not reach basal levels of expression even at 500 nM biotin (**Figure 7b**, **Table 3**). These data were

consistent with the reduced affinity of *Sa*BPL-D200E for the *bioY* oligonucleotide measured in the EMSA analysis. Together these data revealed that the dimerization-impaired mutant, *Sa*BPL D200E, had reduced affinity for binding DNA *in vitro* and transcriptional repressor activity *in vivo*. This suggested the mechanism of resistance to **3** in this *S. aureus* strain is through the depression of biotin synthesis and transport.

Discussion

BPL has been proposed as a drug target for the development of novel anti-infective therapeutics against antibiotic resistant bacteria. Here we utilised a previously reported potent and selective BPL inhibitor (**3**) to study the downstream effects of BPL inhibition in *S. aureus*. We also explored the potential resistance mechanisms to **3** and characterised a new resistance mechanism involving a missense mutation in BPL. The mechanism of action of **3**, and other BPL inhibitors, has previously been proposed to act through the inhibition of the enzymatic function of BPL and reduced protein biotinylation (40). This was determined as a reduction in the antibacterial effect of **3** was observed when BPL was over-expressed or when endogenous biotin was added in the growth media. The current study addressed whether the antimicrobial activity of **3** was also associated with a co-repressor activity. Both biotin and **3** demonstrated similar DNA binding affinities for the *bioO* and *bioY* promoters via *in vitro* EMSA analysis and induced similar kinetics of repression of both *bioY* and bioD (under control of the *bioO* promoter) in *S. aureus*, as measured by qRT-PCR. The data presented here provide evidence that BPL inhibitors not only exert their antibacterial effect by disrupting catalytic activity to prevent protein biotinylation but are also capable of acting as a co-repressor to repress biotin synthesis and biotin transport. This mechanism results in starving the bacteria of an essential micronutrient and, furthermore, helps to reduce the amount of intracellular biotin competing with the inhibitor for binding to BPL. Similar mechanisms whereby antibiotics cause repression of gene expression have previously been reported. One example is the antibiotic pyrithiamine that prevents expression of the essential thiamine biosynthesis pathway by binding in place of thiamine to a riboswitch protein (52). However this is the first evidence of such an effect with a BPL inhibitor.

One of the target genes that *Sa*BPL represses is *bioY* that encodes the substrate specific subunit of the biotin transporter complex responsible for importing biotin into bacteria. Under high biotin concentrations, *Sa*BPL facilitates the repression of the transporter to prevent biotin accumulation when it is not required (22). Here, we demonstrate **3** can function as a co-repressor and, hence, down-regulate expression of the biotin transporter. The mechanism of entry of the biotin-based BPL inhibitors into *S. aureus* has not been established. One possible mechanism is that they are recognised and internalised by BioY through their common biotinyl moiety. Hence, targeting BPL with an inhibitor that acts as a co-repressor will reduce *bioY* expression, limiting the entry of the inhibitor into cells. The mechanism of inhibitor entry must therefore be elucidated to determine whether their action as a co-repressor also reduces their antibacterial efficacy. However, several lines of unpublished data suggest that the BPL

inhibitors are more likely to enter via a different mechanism, potentially passive diffusion or via aquaporins. Indeed no change in the antibacterial activity of **3** was observed when tested against a *bioY* knockout *S. aureus* strain. Similarly, increasing the concentration of exogenous biotin in the growth media did not impede internalisation of a fluorescent BPL inhibitor (29). This suggests BPL inhibitors can still enter *S. aureus* even though they act as co-repressors and down-regulate biotin import.

The ability for *S. aureus* to develop resistance to BPL inhibitors was also analysed. The incidence rate for the generation of spontaneous resistance was low, with no resistant colonies growing after 24 hours, resulting in a resistance rate of less than 1×10^{-9}. Similar low spontaneous resistance rates have been reported for a BPL inhibitor active against *M. tuberculosis* (36). This highlights BPL as a promising antibacterial target with low potential for resistance to develop. Serial passaging of *S. aureus* in sub-optimal concentrations of **3** for 18 days did result in the development of several strains with 4- to 32-fold increased MICs over the starting parent strain. Sequencing of the *bpl* genes revealed a D200E missense mutation present in the most resistant strain. Biochemical characterization of this mutation revealed disrupted homodimerization of *Sa*BPL consequently leading to weakened DNA binding and reduced repressor activity. This residue, D200, is located on a loop within the dimer interface that becomes ordered upon ligand binding. The side chain of D200 forms hydrogen bonding interactions with R122 and hydrophobic interactions with F123 on the opposing subunit (**Supplementary Figure S7c**). By inhibiting dimer formation, D200E disrupted the DNA binding activity of *Sa*BPL and derepression of *Sa*BPL target genes. The mechanism of resistance studies helps support the proposal that the corepressor function of **3** contributes to the antibacterial activity of the compound. Furthermore, disruption of the repressor activity provides an opportunity to resist the action of **3** by increasing the availability of intracellular biotin. The D200E mutant with reduced ability to repress *bioO* and *bioY* would result in increased *de novo* biotin synthesis and biotin transport, respectively. The increased accumulation of intracellular biotin could compete against **3** for binding within the active site of *Sa*BPL. Indeed, increasing exogenous biotin concentrations decreased the antibacterial efficacy of **3**. In support, *S. aureus* can stimulate import and sequester biotin inside the cell in response to increased exogenous biotin concentrations (24). The resistance mechanism *Sa*BPL develops to deregulate its target genes has also been observed in other antibiotic targets in the literature. The antibiotic pyrithiamine binds a riboswitch protein and causes the repression of thiamine biosynthesis genes, leading to thiamine starvation. However, resistance to pyrithiamine is generated through mutations in the binding site of the riboswitch

protein such that the antibiotic can no longer bind and cause gene repression, leading to a derepression of thiamine synthesis (52,53).

The oligomeric state of the *Sa*BPL D200E variant observed here is consistent with previous studies on both *S. aureus* and *E. coli* BPLs, which indicated that modification of residues in the dimerization interface affect repressor ability. The mutations, F123G in *Sa*BPL (24,25) and D197Y (equivalent of D200 in *S. aureus*) in *E. coli* (54,55) demonstrate impaired homodimerization and loss of repressor activity (24,54,55). The interaction between F123 and D200 appears particularly important in *Sa*BPL, supported by the conservation of D200 across 71 non-redundant *Staphylococcal* BPL sequences (**Supplementary Figure S9**). D200 from one subunit forms hydrophobic interactions with F123 on the other subunit to facilitate dimerization. Therefore, it is reasonable to expect the two mutations (D200E and F123G) to function similarly in binding assays. Both resulting mutant proteins have similar K_M values for biotin that are only 2-fold weaker than the wild-type enzyme (Table 2, *Sa*BPL F123G K_M (biotin): 3.06 µM (25)). Therefore, ligand binding and catalytic activities are not impaired. Both mutants are monomeric, even when ligands are bound, but are still capable of binding DNA and inhibiting the transcription of *bioO* and *bioY* albeit to a weaker extent than wildtype protein (24). Therefore, it is proposed that these mutant strains would require a much higher concentration of biotin before *Sa*BPL can be induced to be a fully active transcriptional repressor. Recent analysis of *Sa*BPL F123G has revealed this is not necessarily the case. *Sa*BPL F123G was utilised as a biochemical tool to mimic the oligomeric state of apo-*Sa*BPL, which is also monomeric, for *in vivo* assays where a truly apo-*Sa*BPL cannot be achieved as biotin is required for bacteria to survive. This mutant demonstrated that monomeric *Sa*BPL is capable of binding DNA by a different pathway that does not require ligand binding or dimerization (24). This is divergent from *E. coli* BPL where ligand binding and dimerization are essential pre-requisite steps for DNA-binding. Therefore, *Sa*BPL is proposed to possess an alternative repression mechanism that may assist the survival of the bacteria in low-biotin niches during infection. The *Sa*BPL mutation generated here (D200E), in a *S. aureus* strain resistant to a BPL inhibitor, supports the *Sa*BPL F123G study, being another *Sa*BPL mutant that mimics the oligomeric state of apo-*Sa*BPL but is still active to bind DNA and serve as a repressor *in vivo,* but to a weaker extent. This may allow the resistant *S. aureus* a fine balance between increasing intracellular biotin concentrations to promote a competitive resistance mechanism, whilst limiting biotin synthesis and transport to prevent energetically expensive processes from occurring. Further analysis of *Sa*BPL D200E should be considered

to determine if it does utilise this alternate DNA binding pathway, and the biological relevance for monomeric, apo-*Sa*BPL to bind DNA.

BPL inhibitors have been revealed to exert their antibacterial efficacy through inhibiting the catalytic biotinylation activity of the enzyme and by promoting the repressor activity of the BPL. The BPL inhibitor **3** was revealed to function as a co-repressor to prevent biotin synthesis and import, thus starving the bacteria of an essential micronutrient and preventing competition between biotin and the inhibitor for binding to the target. *S. aureus* strains resistant to **3** were generated, with one of the strains containing a missense mutation in its BPL. This mutation, D200E, provided a resistance mechanism that ultimately circumvented the co-repressor activity of the inhibitor to increase intracellular biotin concentrations to create competition for inhibitor binding to *Sa*BPL. This demonstrates the importance of the co-repressor activity of BPL inhibitors towards their efficacy. Further development of this series of compounds, maintaining both mechanisms of action, will hopefully lead to a valid pre-clinical antibiotic candidate. The low spontaneous resistance rate observed also promotes BPL as a bona fide antibacterial drug target.

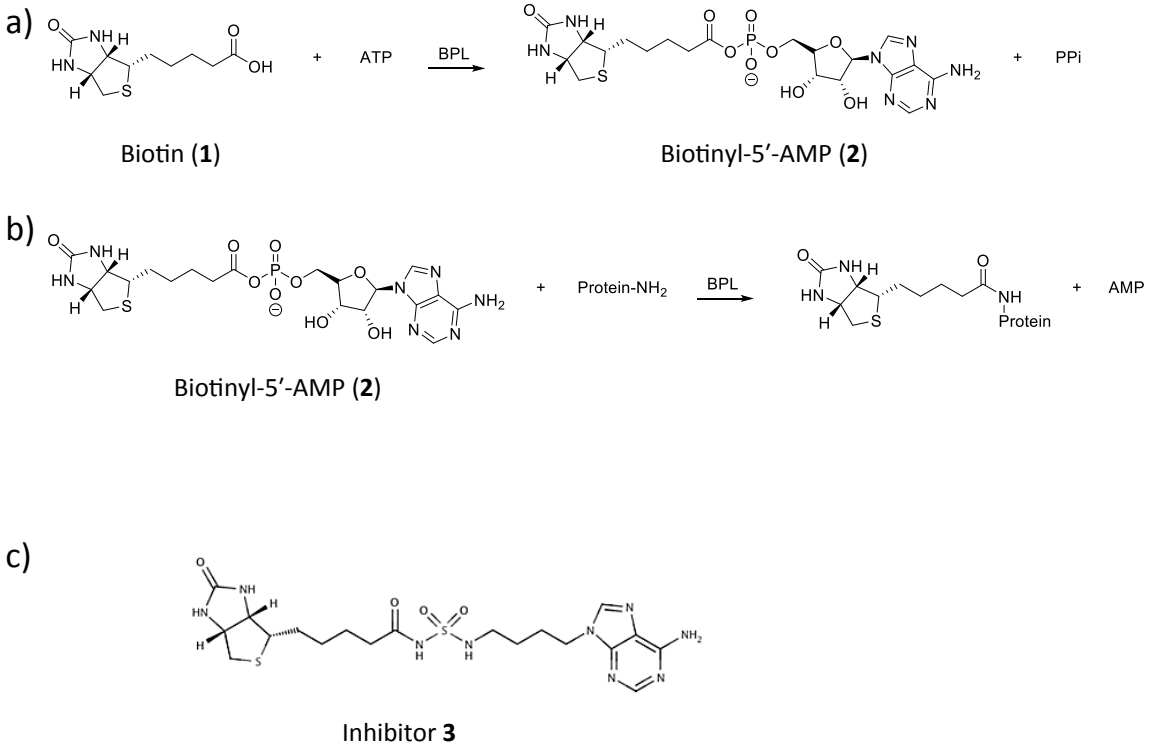


Figure 1. Protein biotinylation catalysed by biotin protein ligase (BPL). a) Biotin (**1**) ligates with ATP to produce the reaction intermediate, biotinyl-5′-AMP (**2**). b.) Biotinyl-5′-AMP is employed to covalently attach the biotinyl moiety onto a specific lysine residue present in the biotin-accepting protein substrate. c.) The chemical structure of BPL inhibitor **3**, a chemical analogue of **2**, is shown.

Table 1: Antibacterial Susceptibity and Cytotoxicity of 3

Species	**MIC (μg/ml)**
Staphylococcus aureus	
Methicillin sensitive (*n*=8)	0.25 – 0.5
Methicillin-resistant (*n*=9)	0.25 – 0.5
Coagulase negative staphylococci (*n*=7)	0.125 – 0.5
Mycobacterium tuberculosis (*n*=1)	55
Streptococcus pneumoniae (*n*=6)	>32
Enterococococcus faecalis (*n*=3)	>128
Enterococococcus faecalis (*n*=3)	>128
Enterococcus faecium (*n*=5)	>128
Escherichia coli (*n*=1)	0.25 – 0.5
Cell lines	**EC_{50} (μg/ml)**
HepG2	>250
HEK293	>250

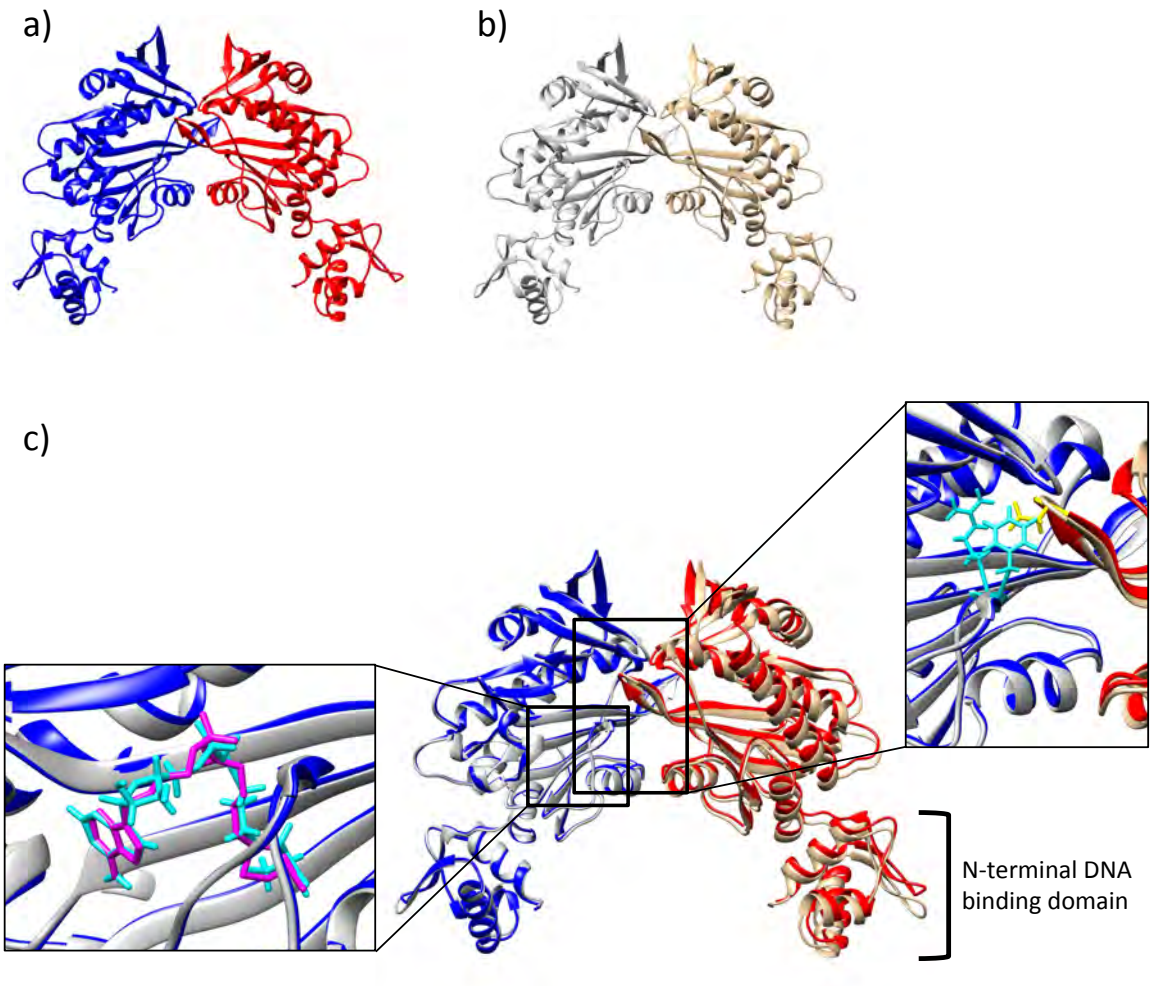


Figure 2: *Sa*BPL structures

Structures of the *Sa*BPL homodimer in complex with a.) biotinyl-5′-AMP (blue and red) (PDB: 3RIR) (56) and b.) inhibitor **3** (grey and tan) (PDB: 6ORU) (40). c.) An overlay of these two structures, with insert panels showing the conserved binding mode of biotinyl-5′-AMP (cyan) and **3** (purple), and the residues involved in dimerization including R122 and F123 (cyan) and D200 (yellow) from the opposite subunit.

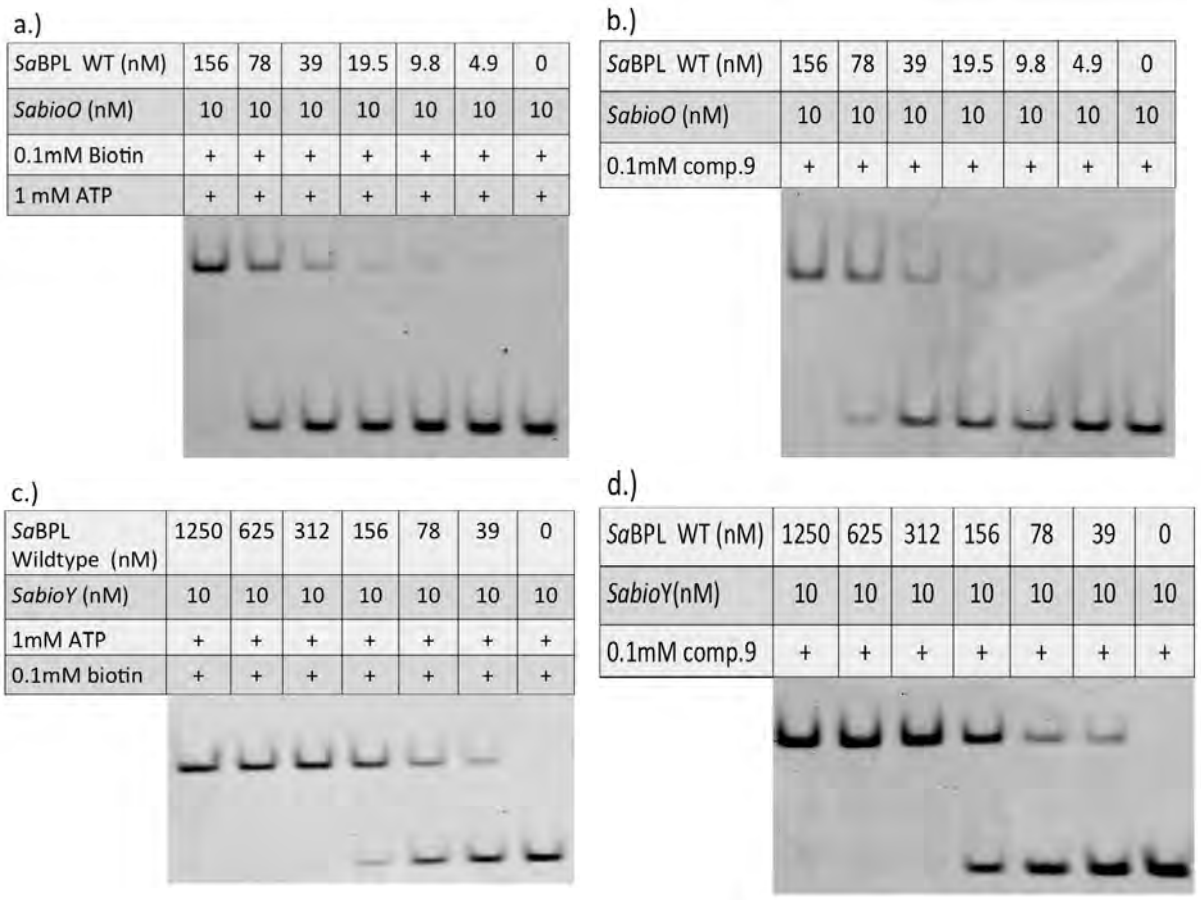


Figure 3: *Sa*BPL in complex with 3 is competent to bind DNA. *Sa*BPL binding to the *SabioO* oligonucleotide in the presence of a.) biotinyl-5′-AMP and b.) 3, and the *SabioY* oligonucleotide in the presence of c.) biotinyl-5′-AMP and d.) **3**. Control reactions for EMSA analysis containing biotinyl-5′-AMP were prepared by adding 0.1 mM biotin, 1 mM ATP and 2.5% DMSO. These were compared to reactions containing 0.1 mM of compound **3** and 2.5% DMSO.

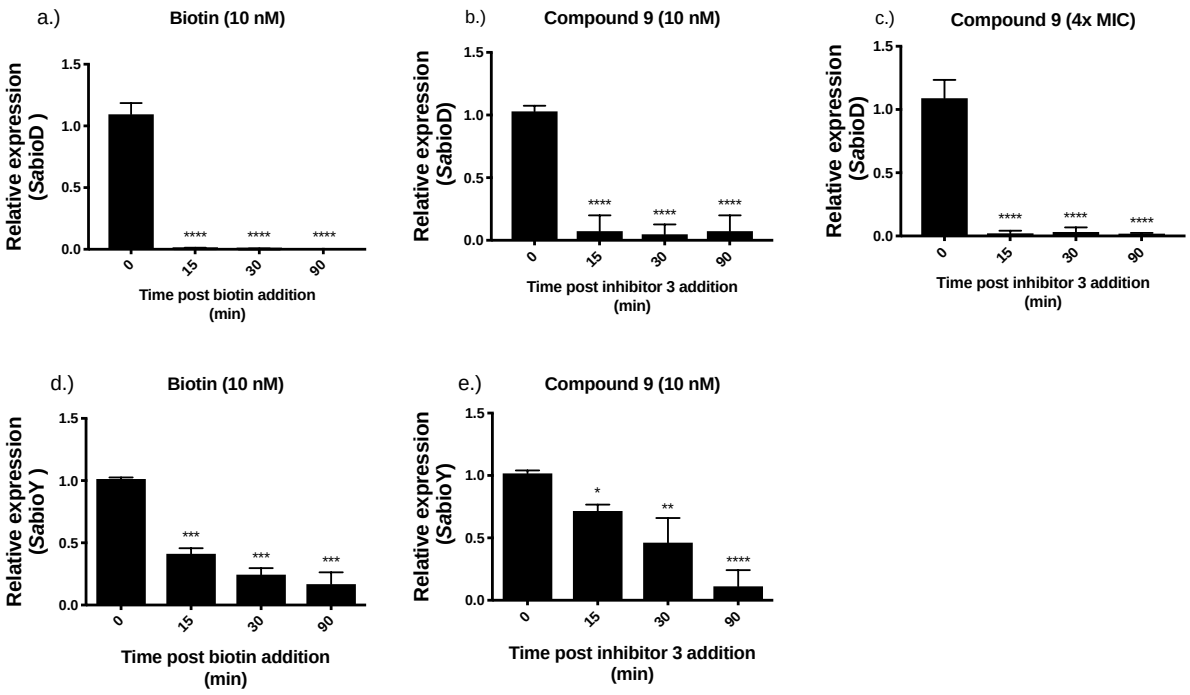


Figure 4: Inhibitor 3 induces the repression of *bioD* and *bioY*. qRT-PCR measured the repression of *SabioD* following treatment with a.) 10 nM biotin b.) 10 nM compound **3** and c.) 3.9 µM (4x MIC) of **3**, and the repression of *SabioY* in response to d.) 10 nM biotin and e.) 10 nM of **3**. Transcript levels were normalized against an internal control (16s rRNA) and calculated relative to time = 0. Error bars represent S.E.M from 3 independent biological replicates (n=3), * = p <0.05, ** = p <0.01, *** = p >0.001, **** = p <0.0001.

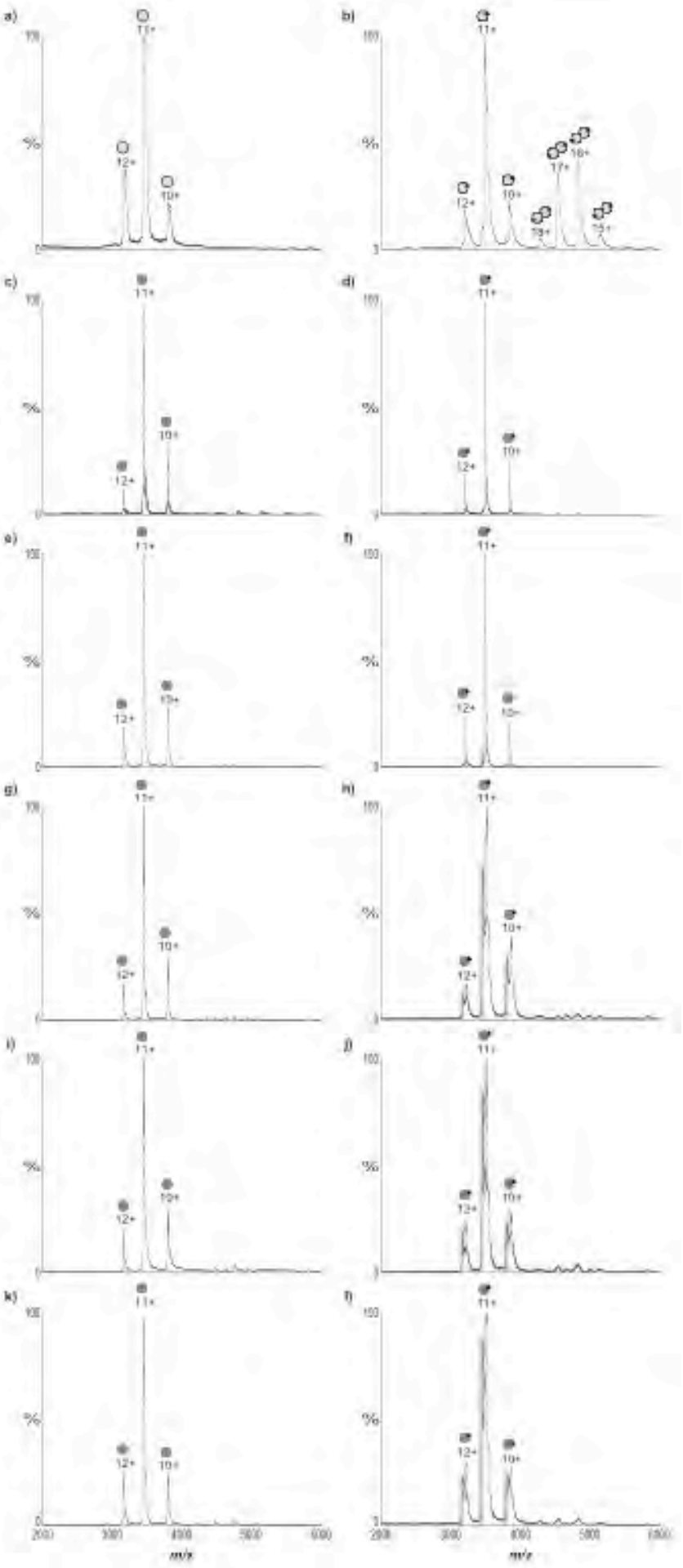

a)
11+
12+
10+
b)
11+
12+
10+
17+
16+
c)
11+
12+
10+
d)
11+
12+
10+
e)
11+
12+
10+
f)
11+
12+
g)
11+
12+
10+
h)
11+
12+
10+
i)
11+
12+
10+
j)
11+
12+
10+
k)
11+
12+
10+
l)
11+
12+
10+
%
m/z

Figure 5. Native MS reveals *Sa*BPL D200E is monomeric in the presence and absence of ligands. nESI-MS to measure the oligomeric state of a.) 10 µM apo-wild-type *Sa*BPL (outlined light grey circles), b.) 10 µM holo-wild-type *Sa*BPL, c.) 1.4 µM apo-*Sa*BPL D200E (dark grey circles), d.) 1.4 µM holo-*Sa*BPL D200E, e.) 11.25 µM apo-*Sa*BPL D200E, f.) 11.25 µM holo-*Sa*BPL D200E, g.) 22.5 µM apo-*Sa*BPL D200E, h.) 22.5 µM holo-*Sa*BPL D200E, i.) 45 µM apo-*Sa*BPL D200E, j.) 45 µM holo-*Sa*BPL D200E, k.) 90 µM apo-*Sa*BPL D200E and l.) 90 µM holo-*Sa*BPL D200E. Monomer is represented by a single sphere, dimer by conjoined spheres and the presence of biotinyl-5′-AMP by a black triangle.

Table 2. Comparisons in the kinetic activity and inhibition by 3 of wildtype *Sa*BPL and *Sa*BPL D200E. SEM calculated from independent biological replicates (n=3).

***Sa*BPL**	**K_M (biotin) (μM)**	**K_i 3 (μM)**
Wild-type	1.8 ± 0.3	0.005 ± 0.002
D200E	3.8 ± 0.4	0.011 ± 0.004

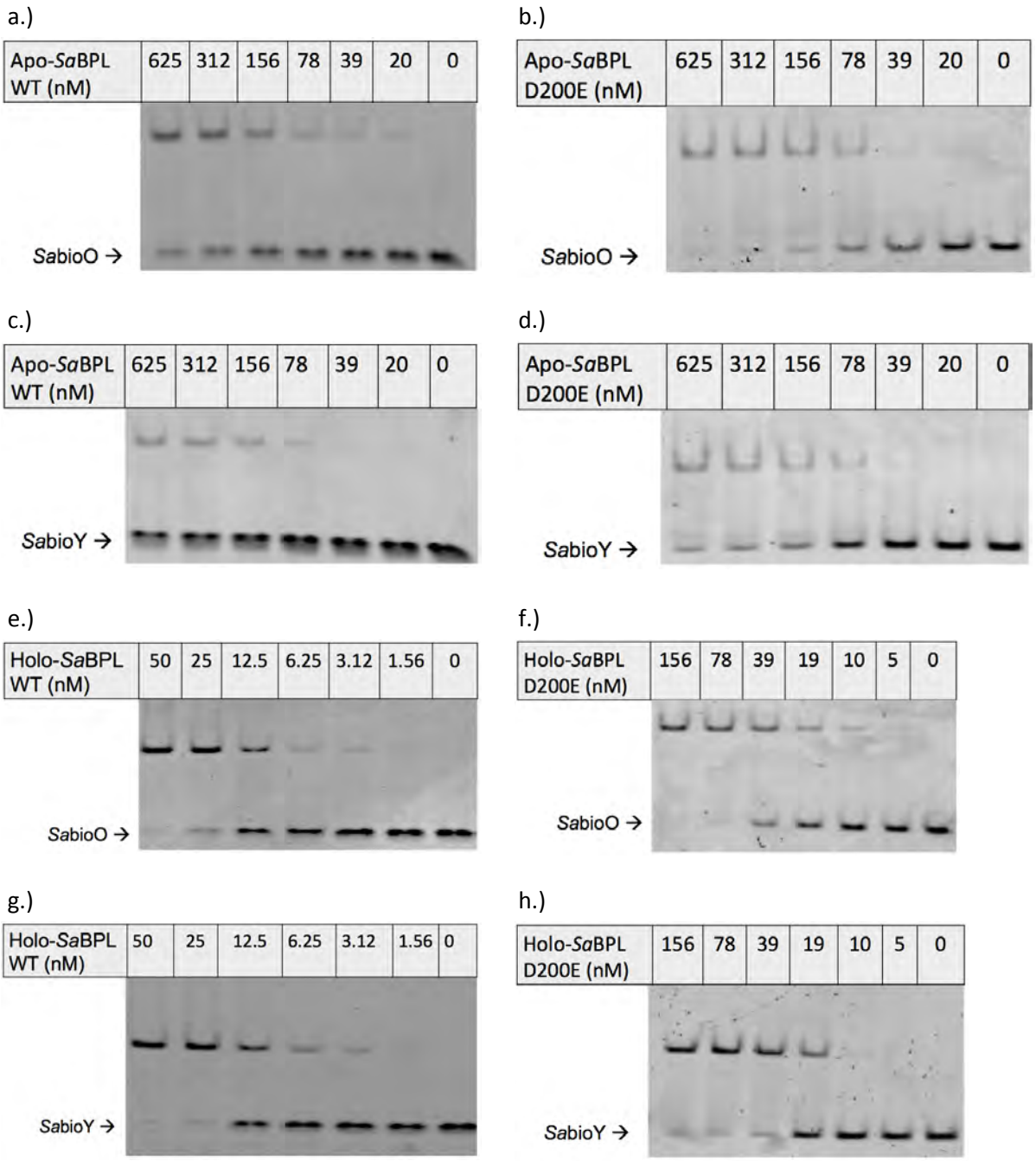


Figure 6: *Sa*BPL D200E has reduced DNA binding activity compared to wild-type *Sa*BPL. EMSA analysis to measure the binding of a.) apo-*Sa*BPL wild-type binding to *bioO*, b.) apo-*Sa*BPL D200E binding to *bioO*, c.) apo-*Sa*BPL wild-type binding to *bioY* and d.) apo-*Sa*BPL D200E binding to *bioY*, e.) holo-*Sa*BPL wild-type binding to *bioO*, f.) holo-*Sa*BPL D200E binding to *bioO*, g.) holo-*Sa*BPL wild-type binding to *bioY* and h.) holo-*Sa*BPL D200E binding to *bioY*.

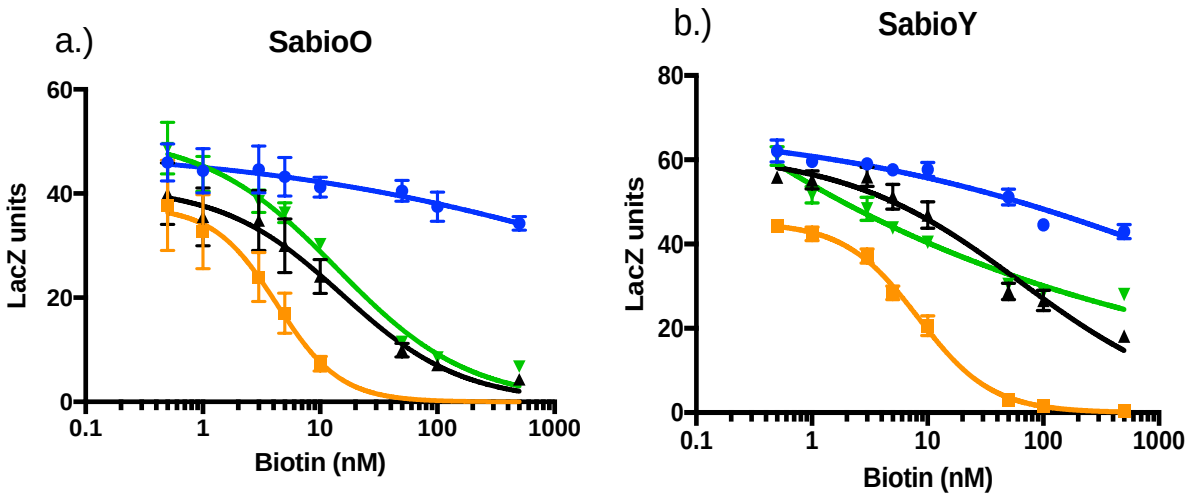


Figure 7: Biotin-regulated repressor activity of *Sa*BPL and mutants. *In vivo* assay was performed in an *E. coli* reporter strain containing chromosomally integrated repressor and promoter constructs. *LacZ* units were calculated by subtracting the value generated by the control strain containing no integrated promoter construct (*LacZ* unit ≤ 10). Expression of the β-lactamase reporter gene, under the control of promoters a.) *bioO* or b.) *bioY*, was measured in media containing varying concentrations of biotin. Regulators *Sa*BPL (orange), *Sa*BPL D200E (green) and *Sa*BPL F123G (black) were analysed alongside a control strain that harboured no repressor protein (blue). Error bars represent S.E.M from independent biological replicates (n = 6).

Table 3: $K_{R\ biotin}$ values revealing the repressor activity of wild-type *Sa*BPL and *Sa*BPL D200E in the *in vivo* β-galactosidase assay. The amount of biotin to reach half-maximum repression ($K_{R\ biotin}$) was calculated from the results generated in **Figure 7** using GraphPad Prism. The significance of the difference in repression of the *SabioO* promoter between the two enzymes was determined using a two tailed t-test. (n=6) * $p < 0.05$.

***Sa*BPL**	**Promoter**	**Half maximum repression $K_{R\ biotin}$ (nM)**
Wild type	*Sa*bioO	4.3 ± 1.9
D200E	*Sa*bioO	13.9 ± 3.4
Wild type	*Sa*bioY	8.2 ± 0.7
D200E	*Sa*bioY	≥500

References:

1. Cooper, A., Shlaes, D. (2011) Fix the antibiotics pipeline. *Nature* **472**, 32
2. Butler, M. S., Blaskovich, M. A., and Cooper, M. A. (2016) Antibiotics in the clinical pipeline at the end of 2015. *J. Antibiot. (Tokyo)*, 1-22
3. (2014) World Health Organisation, Antimicrobial resistance: Global report on resistance.
4. (2015) World Health Organisation: Global Action Plan on Antimicrobial Resistance.
5. Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I., and Miller, A. A. (2015) ESKAPEing the labyrinth of antibacterial discovery. *Nat. Rev. Drug Discov.* **14**, 529-542
6. Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **6**, 29-40
7. Feng, J., Paparella, A. S., Booker, G. W., Polyak, S. W., and Abell, A. D. (2016) Biotin protein ligase is a target for new antibacterials. *Antibiotics (Basel)* **5**
8. Paparella, A. S., Soares da Costa, T. P., Yap, M. Y., Tieu, W., Wilce, M. C. J., Booker, G. W., Abell, A. D., and Polyak, S. W. (2014) Structure guided design of Biotin protein ligase inhibitors for antibiotic discovery. *Curr. Top. Med. Chem.* **14**, 4-20
9. Sternicki, L. M., Wegener, K. L., Bruning, J. B., Booker, G. W., and Polyak, S. W. (2017) Mechanisms governing precise protein biotinylation. *Trends. Biochem. Sci.* **42**, 383-394
10. 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., and Charles, I. G. (2009) Comprehensive identification of essential Staphylococcus aureus genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **10**, 291
11. Forsyth, R. A., Haselbeck, R. J., Ohlsen, K. L., Yamamoto, R. T., Xu, H., Trawick, J. D., Wall, D., Wang, L., Brown-Driver, V., Froelich, J. M., C, K. G., King, P., McCarthy, M., Malone, C., Misiner, B., Robbins, D., Tan, Z., Zhu Zy, Z. Y., Carr, G., Mosca, D. A., Zamudio, C., Foulkes, J. G., and Zyskind, J. W. (2002) A genome-wide strategy for the identification of essential genes in Staphylococcus aureus. *Mol Microbiol* **43**, 1387-1400
12. Polyak, S. W., Abell, A. D., Wilce, M. C. J., Zhang, L., and Booker, G. W. (2012) Structure, function and selective inhibition of bacterial acetyl-CoA carboxylase. *Appl. Microbiol. Biotechnol.* **93**, 983-992
13. Yao, J., and Rock, C. O. (2017) Bacterial fatty acid metabolism in modern antibiotic discovery. *Biochim Biophys Acta Mol Cell Biol Lipids* **1862**, 1300-1309
14. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Pendini, N. R., Polyak, S. W., Pedersen, D. S., Morona, R., Turnidge, J. D., Wallace, J. C., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2012) Selective inhibition of Biotin protein ligase from *Staphylococcus aureus*. *J. Biol. Chem.* **287**, 17823-17832
15. Duckworth, B. P., Geders, T. W., Tiwari, D., Boshoff, H. I., Sibbald, P. A., Barry, C. E., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2011) Bisubstrate adenylation inhibitors of Biotin protein ligase from *Mycobacterium tuberculosis*. *Chem. Biol.* **18**, 1432-1441
16. Shi, C., Tiwari, D., Wilson, D. J., Seiler, C. L., Schnappinger, D., and Aldrich, C. C. (2013) Bisubstrate Inhibitors of Biotin Protein Ligase in Mycobacterium tuberculosis Resistant to Cyclonucleoside Formation. *ACS Med Chem Lett* **4**
17. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Zvarec, O., Bell, J. M., Turnidge, J. D., Wallace, J. C., Booker, G. W., Wilce, M. C., Abell, A. D., and Polyak, S. W. (2012)

Biotin analogues with antibacterial activity are potent inhibitors of biotin protein ligase. *ACS Med Chem Lett* **3**, 509-514
18. Feng, J., Paparella, A. S., Tieu, W., Heim, D., Clark, S., Hayes, A., Booker, G. W., Polyak, S. W., and Abell, A. D. (2016) New series of BPL inhibitors to probe the ribose-binding pocket of *Staphylococcus aureus* Biotin protein ligase. *ACS Med. Chem. Lett.* **7**, 1068-1072
19. Benton, B. M., Zhang, J. P., Bond, S., Pope, C., Christian, T., Lee, L., Winterberg, K. M., Schmid, M. B., and Buysse, J. M. (2004) Large-scale identification of genes required for full virulence of Staphylococcus aureus. *J Bacteriol* **186**, 8478-8489
20. Bae, T., Banger, A. K., Wallace, A., Glass, E. M., Aslund, F., Schneewind, O., and Missiakas, D. M. (2004) Staphylococcus aureus virulence genes identified by bursa aurealis mutagenesis and nematode killing. *Proc Natl Acad Sci U S A* **101**, 12312-12317
21. Rodionov, D. A., Mironov, A. A., and Gelfand, M. S. (2002) Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. *Genome Res* **12**, 1507-1516
22. Satiaputra, J., Eijkelkamp, B. A., McDevitt, C. A., Shearwin, K. E., Booker, G. W., and Polyak, S. W. (2018) Biotin-mediated growth and gene expression in *Staphylococcus aureus* is highly responsive to environmental biotin. *Appl Microbiol Biotechnol* **102**, 3793-3803
23. Satiaputra, J., Shearwin, K. E., Booker, G. W., and Polyak, S. W. (2016) Mechanisms of biotin-regulated gene expression in microbes. *Synth Syst Biotechnol* **1**, 17-24
24. Satiaputra, J., Sternicki, L. M., Hayes, A. J., Pukala, T. L., Booker, G. W., Shearwin, K. E., and Polyak, S. W. (2019) Native mass spectrometry identifies an alternative DNA-binding pathway for BirA from Staphylococcus aureus. *Sci Rep* **9**, 2767
25. Soares da Costa, T. P., Yap, M. Y., Perugini, M. A., Wallace, J. C., Abell, A. D., Wilce, M. C. J., Polyak, S. W., and Booker, G. W. (2014) Dual roles of F123 in protein homodimerization and inhibitor binding to biotin protein ligase from *Staphylococcus aureus*. *Mol. Microbiol.* **91**, 110-120
26. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterization of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci.* **22**, 762-773
27. Brown, P. H., Cronan, J. E., Grøtli, M., and Beckett, D. (2004) The biotin repressor: modulation of allostery by corepressor analogs. *J. Mol. Biol.* **337**, 857-869
28. Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D., and Matthews, B. W. (2006) Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution. *J Mol Biol* **357**, 509-523
29. Paparella, A. S., Lee, K. J., Hayes, A. J., Feng, J., Feng, Z., Cini, D., Deshmukh, S., Booker, G. W., Wilce, M. C. J., Polyak, S. W., and Abell, A. D. (2018) Halogenation of Biotin Protein Ligase Inhibitors Improves Whole Cell Activity against Staphylococcus aureus. *ACS Infect Dis* **4**, 175-184
30. Feng, J., Paparella, A. S., Tieu, W., Heim, D., Clark, S., Hayes, A., Booker, G. W., Polyak, S. W., and Abell, A. D. (2016) New Series of BPL Inhibitors To Probe the Ribose-Binding Pocket of *Staphylococcus aureus* Biotin Proteing Ligase *ACS Med Chem Lett* **7**, 1068-1072
31. Tieu, W., Jarrad, A. M., Paparella, A. S., Keeling, K. A., Soares da Costa, T. P., Wallace, J. C., Booker, G. W., Polyak, S. W., and Abell, A. D. (2014) Heterocyclic acyl-phosphate bioisostere-based inhibitors of *Staphylococcus aureus* biotin protein ligase. *Bioorg Med Chem Lett* **24**, 4689-4693
32. Tieu, W., Polyak, S. W., Paparella, A. S., Yap, M. Y., Soares da Costa, T. P., Ng, B., Wang, G., Lumb, R., Bell, J. M., Turnidge, J. D., Wilce, M. C. J., Booker, G. W., and

Abell, A. D. (2015) Improved Synthesis of Biotinol-5′-AMP: Implications for Antibacterial Discovery. *ACS Med Chem Lett* **6**, 216-220
33. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Zvarec, O., Bell, J. M., Turnidge, J. D., Wallace, J. C., Booker, G. W., Wilce, M. C. J., Abell, A. D., and Polyak, S. W. (2012) Biotin Analogues with Antibacterial Activity Are Potent Inhibitors of Biotin Protein Ligase. *ACS Med Chem Lett* **3**, 509-514
34. Soares da Costa, T. P., Tieu, W., Yap, M. Y., Pendini, N. R., Polyak, S. W., Pedersen, D. S., Morona, R., Turnidge, J. D., Wallace, J. C., Wilce, M. C. J., Booker, G. W., and Abell, A. D. (2012) Selective inhibition of Biotin Protein Ligase from *Staphylococcus aureus*. *J Biol Chem* **287**, 17823-17832
35. Tiwari, D., Park, S. W., Essawy, M. M., Dawadi, S., Mason, A., Nandakumar, M., Zimmerman, M., Mina, M., Pin Ho, H., Engelhart, C. A., Ioerger, T., Sacchettini, J. C., Rhee, K., Ehrt, S., Aldrich, C. C., Dartois, V., and Schnappinger, D. (2018) Targeting protein biotinylation enhances tuberculosis chemotherapy. *Sci Transl Med* **10**
36. Bockman, M. R., Engelhart, C. A., Dawadi, S., Larson, P., Tiwari, D., Ferguson, D. M., Schnappinger, D., and Aldrich, C. C. (2018) Avoiding Antibiotic Inactivation in *Mycobacterium tuberculosis* by Rv3406 through Strategic Nucleoside Modification. *ACS Infect Dis* **4**, 1102-1113
37. Bockman, M. R., Kalinda, A. S., Petrelli, R., De la Mora-Rey, T., Tiwari, D., Liu, F., Dawadi, S., Nandakumar, M., Rhee, K. Y., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2015) Targeting *Mycobacterium tuberculosis* Biotin Protein Ligase (MtBPL) with Nucleoside-Based Bisubstrate Adenylation Inhibitors. *J Med Chem* **58**, 7349-7369
38. Shi, C., Tiwari, D., Wilson, D. J., Seiler, C. L., Schnappinger, D., and Aldrich, C. C. (2013) Bisubstrate Inhibitors of Biotin Protien Ligase in *Mycobacterium tuberculosis* Resistant to Cyclonucleoside Formation. *American Chemical Society Medicinal Chemistry Letters* **4**, 1213-1217
39. Duckworth, B. P., Geders, T. W., Tiwari, D., Boshoff, H. I., Sibbald, P. A., Barry III, C. E., Schnappinger, D., Finzel, B. C., and Aldrich, C. C. (2011) Bisubstrate Adenylation Inhibitors of Biotin Protein Ligase from *Mycobacterium tuberculosis*. *Chem Biol* **18**, 1432-1441
40. Lee, K. J. (2019) Sulfonamide-based inhibitors of biotin protein ligase as new antibiotic leads.
41. Hellman, L. M., and Fried, M. G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc* **2**, 1849-1861
42. Pope, C. F., O'Sullivan, D. M., McHugh, T. D., and Gillespie, S. H. (2008) A practical guide to measuring mutation rates in antibiotic resistance. *Antimicrob Agents Chemother* **52**, 1209-1214
43. Luria, S. E., and Delbrück, M. (1943) Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **28**, 491-511
44. Young, K. (2006) In vitro antibacterial resistance selection and quantitation. *Curr Protoc Pharmacol* **Chapter 13**, Unit13A.16
45. Friedman, L., Alder, J. D., and Silverman, J. A. (2006) Genetic changes that correlate with reduced susceptibility to daptomycin in Staphylococcus aureus. *Antimicrob Agents Chemother* **50**, 2137-2145
46. Brune, I., Gotker, S., Schneider, J., Rodionov, D. A. and Tauch, A. (2012) Negative transcriptional control of biotin metabolism genes by the TetR-type regulator BioQ in biotin-auxotrophic Corynebacterium glutamicum ATCC 13032. *J Biotechnol* **159**, 225-234
47. Pendini, N. R., Polyak, S. W., Booker, G. W., Wallace, J. C., and Wilce, M. C. J. (2008) Purification, crystallization and preliminary crystallographic analysis of biotin

protein ligase from *Staphylococcus aureus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64**, 520-523
48. Chapman-Smith, A., Turner, D. L., Cronan, J. E., Jr., Morris, T. W., and Wallace, J. C. (1994) Expression, biotinylation and purification of a biotin-domain peptide from the biotin carboxy carrier protein of *Escherichia coli* acetyl-CoA carboxylase. *Biochem J* **302 (Pt 3)**, 881-887
49. St-Pierre, F., Cui, L., Priest, D. G., Endy, D., Dodd, I. B., and Shearwin, K. E. (2013) One-step cloning and chromosomal integration of DNA. *ACS Synthet. Biol.* **2**, 537-541
50. Garcia, L. S. (2010) *Clinical microbiology procedures handbook*, American Society for Microbiology Press
51. Odds, F. C. (2003) Synergy, antagonism, and what the chequerboard puts between them. *Journal of Antimicrobial Chemotherapy* **52**, 1-1
52. Blount, K. F., and Breaker, R. R. (2006) Riboswitches as antibacterial drug targets. *Nat Biotechnol* **24**, 1558-1564
53. Pedrolli, D. B., Matern, A., Wang, J., Ester, M., Siedler, K., Breaker, R., and Mack, M. (2012) A highly specialized flavin mononucleotide riboswitch responds differently to similar ligands and confers roseoflavin resistance to Streptomyces davawensis. *Nucleic Acids Res* **40**, 8662-8673
54. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., and Matthews, B. W. (1992) *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc Natl Acad Sci U S A* **89**, 9257-9261
55. Buoncristiani, M. R., Howard, P. K., and Otsuka, A. J. (1986) DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (BirA) of *Escherichia coli*. *Gene* **44**, 255-261
56. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterisation of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci* **22**, 762-773

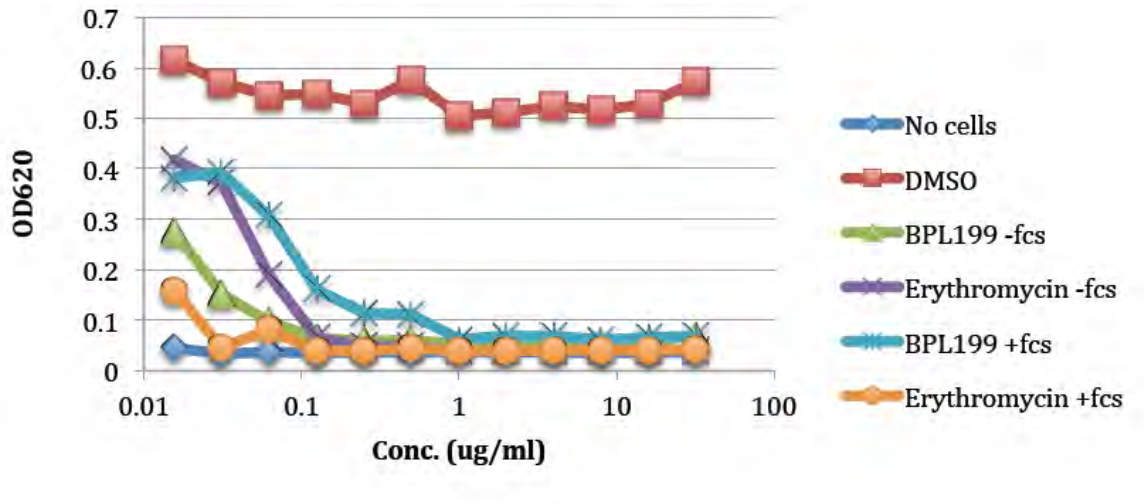
SUPPLEMENTARY INFORMATION

Probing the mechanisms of action and resistance of BPL inhibitors in *Staphylococcus aureus* using an antibacterial sulfonyl based mimic of the reaction intermediate, biotinyl-5′-AMP

Hayes, A.J.[1, 3*], Satiaputra, J.[1, 4*], Sternicki, L.M.[1], Paparella, A.S.[1, 5], Feng, Z.[1, 6], Rodriguez, B.B.[2], Feng, J. [2], Tieu, W.[2, 7], Eijkelkamp, B.A.[1], Shearwin, K.[1], Pukala, T.L.[2], Abell, A. D.[2], Booker, G.W.[1] & Polyak, S.W.[1, 8]

[1] School of Biological Sciences, University of Adelaide, South Australia 5005, Australia
[2] School of Physical Sciences, University of Adelaide, South Australia 5005, Australia
[3] Present address: School of Biomedical Sciences, The Peter Doherty Institute for Infection and Immunity, The University of Melbourne and The Royal Melbourne Hospital, Melbourne, Victoria 3000, Australia.
[4] Present address: Faculty of Health and Medical Sciences, UWA Centre for Medical Research, University of Western Australia, Perth, Western Australia, 6009 Australia
[5] Present address: Albert Einstein College of Medicine, New York City, New York, 10461 United States of America
[6] Present address: School of Natural Sciences and School of Medicine, The University of Tasmania, Hobart, Tasmania, 7001 Australia
[7] Present address: South Australian Health and Medical Research Institute, Adelaide, South Australia, 5000 Australia
[8] Present address: School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, South Australia, 5001 Australia

*** These authors contributed equally to this work**



9

Figure S1: Antibacterial efficacy of **3** (also called BPL199) in the presence (blue) and absence (green) of 20% FCS. This is compared to the antibacterial efficacy of erythromycin control in the presence (orange) and absence (purple) of 20% FCS.

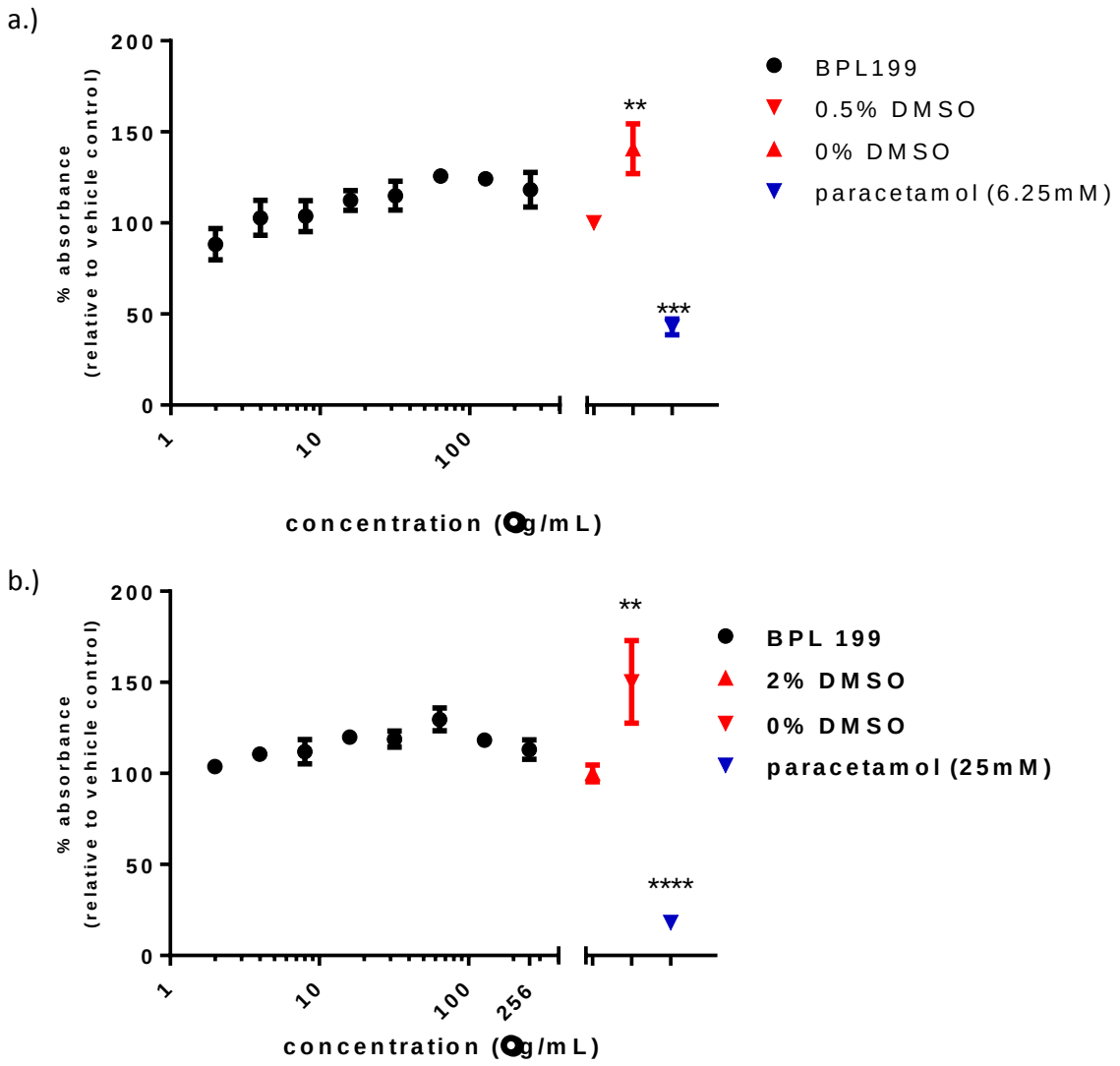


Figure S2: Cytotoxicity of **3** (BPL199) against both a.) HEK293 and b.) HepG2 cell lines with paracetamol as a positive control. The WST-1 reagent was used to detect metabolic activity after 48 hours treatment. No significant toxicity was observed at any concentration. Paracetamol at 6.25mM and 25mM respectively exhibited >60% reduction of growth ($p<0.001$).

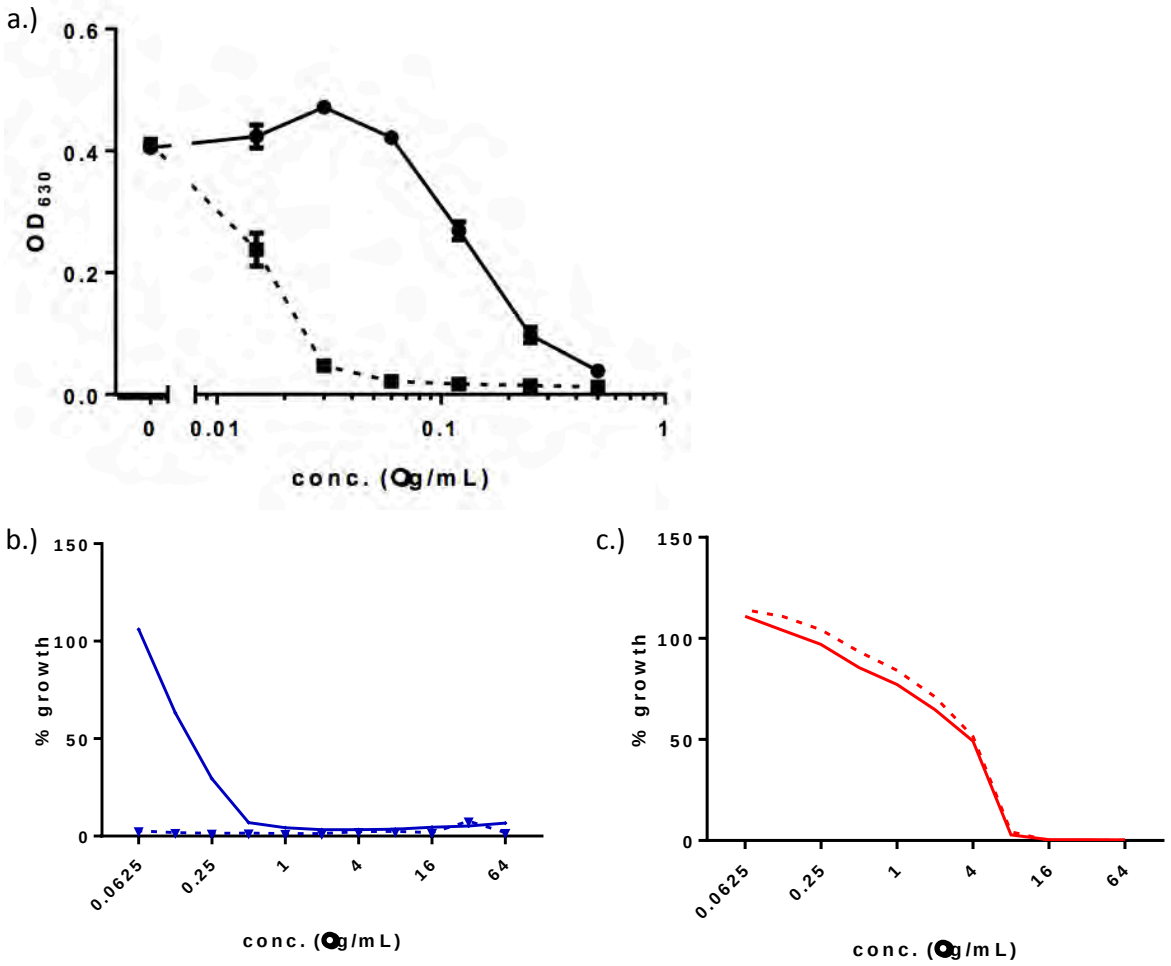


Figure S3: Mechanism of action of compound **3** against *S. aureus***.** Susceptibility of *S. aureus* RN4220 containing pCN51-NcoI control vector (dashed line) and pCN51-BPL overexpression vector (solid line) to a.) inhibitor **3**, b.) literature BPL inhibitor biotinol-5′-AMP and c.) non-BPL targeting antibiotic amoxicillin. Assays were performed in triplicate and normalised to no compound control.

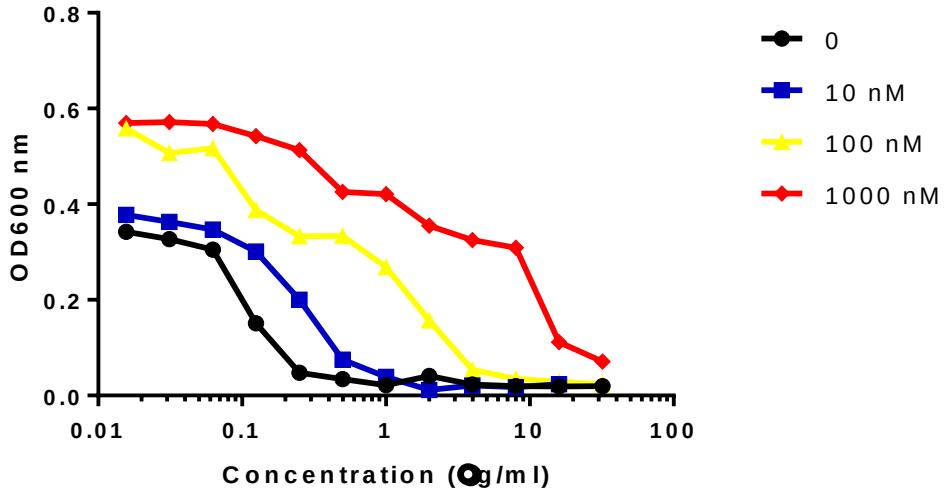


Figure S4: Effect of increased exogenous biotin concentration on antimicrobial effect of **3** against *S. aureus* ATCC 49775. Mueller Hinton broth was supplemented with either no additional biotin (black circles), 10 nM biotin (blue squares), 100 nM biotin (yellow triangles) or 1 µM biotin (red diamonds).

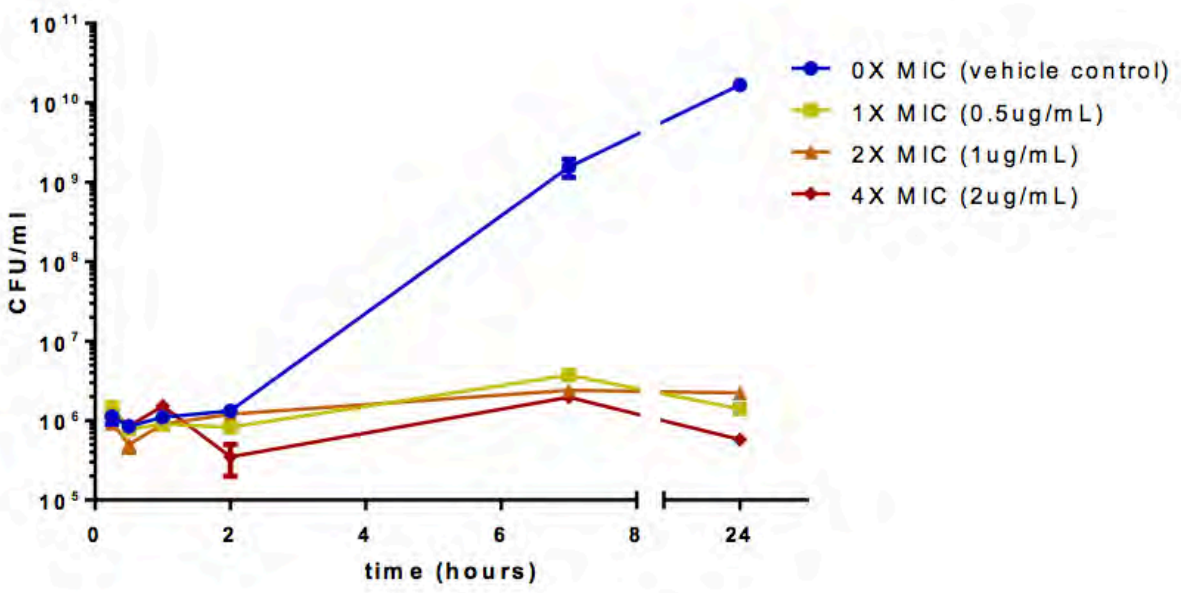
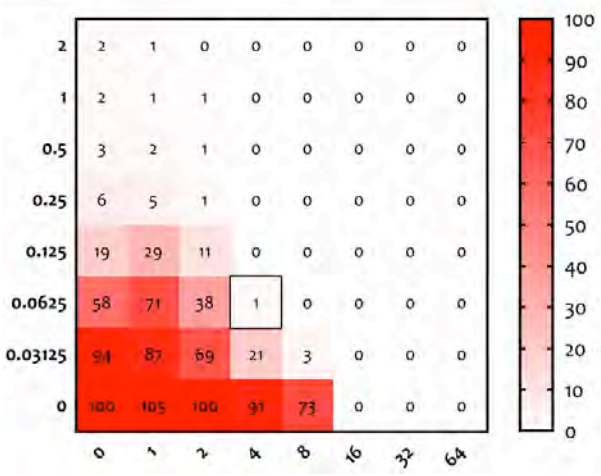
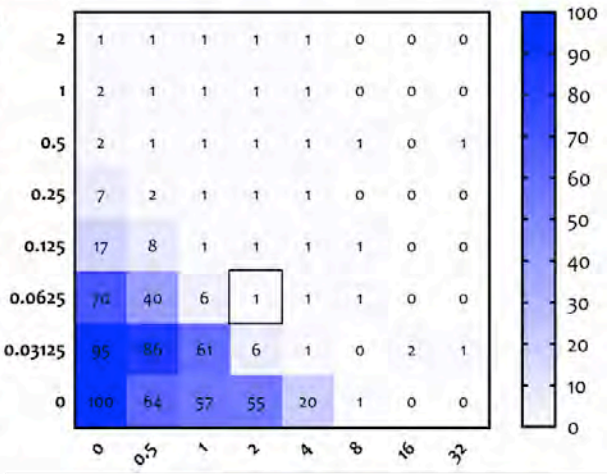


Figure S5: Kinetics of the antibacterial activity of **3**. *S. aureus* ATCC49775 was grown in the presence of no compound (blue), 1× MIC (yellow squares), 2× MIC (orange triangles), 4× MIC (red diamonds) and cell numbers enumerated to determine if *S. aureus* cell death was observed.

a)



b)



c)

Antibacterial agent	Class	Mechanism of action	∑ FIC index with BPL199	Synergy
Synergistic activity				
Methicillin	β-lactam	Cell wall synthesis	0.375	Yes
Streptomycin	Aminoglycosides	Protein synthesis	0.375	Yes
No synergistic activity				
Vancomycin	Glycopeptides	Cell wall synthesis	0.625	No
Daptomycin	Lipopeptides	Cell membrane depolarisation; nucleotide and protein synthesis	0.625	No
Erythromycin	Macrolides	Protein synthesis	0.625	No
Chloramphenicol	Chloramphenicols	Protein synthesis	0.5625	No
Tetracycline	Tetracyclines	Protein synthesis	0.5625	No

Figure S6: Checkerboard assay results of **3** (BPL199) with a) methicillin and b) streptomycin. c.) Antibacterial synergy of **3** with licenced antibiotics. Due to variability in the plate reader a consistent value of ≤2% growth in separate plates (n=3) was taken as optical clarity.

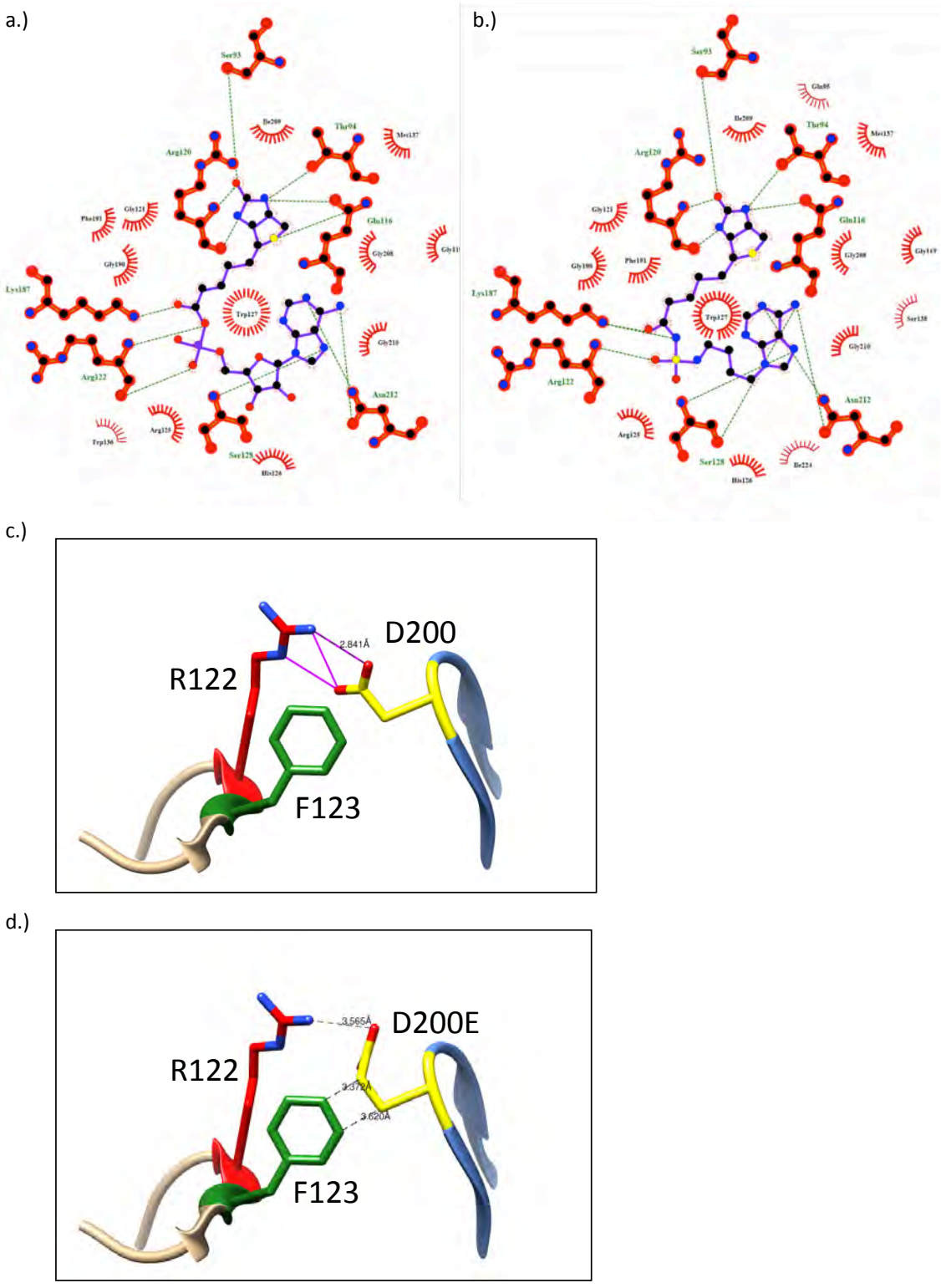


Figure S7. Molecular binding interactions between *Sa*BPL and a.) biotinyl-5′-AMP (PDB: 3RIR) (1) and b.) compound **3** (PDB: 6ORU) (2) determined by ligplot+ (3,4). c.) and d.) Modelling the effect of the D200E mutation (yellow) in a *Sa*BPL crystal structure (PDB:3RIR) (using UCSF Chimera swapaa tool, and assuming no movement in adjacent atoms). c.) In

wild-type *Sa*BPL D200 forms hydrogen bonds with R122 (red) and hydrophobic interactions with F123 of the opposing subunit in the dimer interface. d.) When D200 is mutated to E, there is a loss of these interactions as the distance between these residues increases to prevent steric clashes.

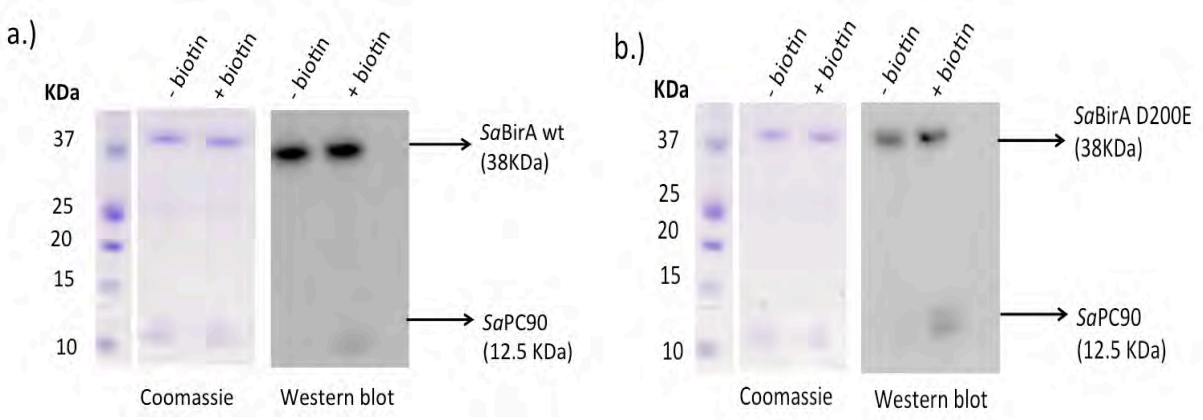


Figure S8. Both a.) wild-type *Sa*BPL and b.) *Sa*BPL D200E were purified in the apo state, with no detectable reaction intermediate bound. Western blots were detected using Alexa 488 conjugated to streptavidin to measure biotinylated *Sa*PC90 (biotin domain) after incubation with *Sa*BPL in the presence and absence of biotin. Coomassie stained gel was used as a loading control.

*A small proportion of biotin might be covalently linked to a non-specific lysine residue within *Sa*BPL during purification, which results in detection of *Sa*BirA in Western blot.

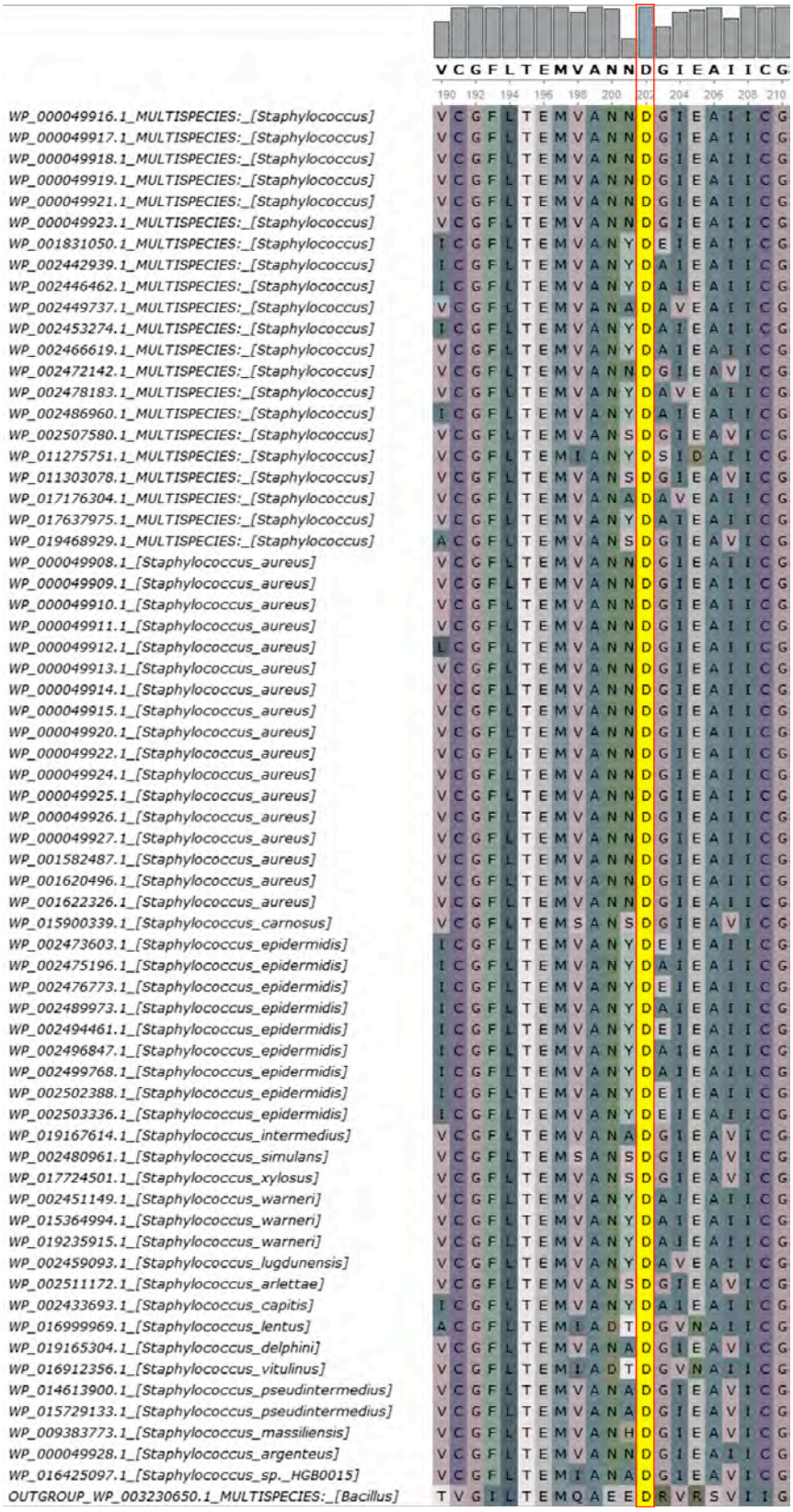


Figure S9. Conservation of the D200 residue across the 71 non-redundant *Staphylococcal* BPL sequences from the protein cluster PCLA_885364. The outgroup *Bacillus subtillus* was also included. Sequences were aligned with Clustal omega (5) and viewed using UGENE (6).

Supplementary Table 1. Molecular weights of complexes measured by native nano-electrospray ionisation mass spectrometry. Masses are calculated based on the loss of the initiating methionine.

***Sa*BirA Sample**	**Measured MW (Da)**	**Complex components**	**Calculated MW (Da)**
Apo-wild type *Sa*BirA 10 µM	37892	Monomer	37892
Holo-wild type *Sa*BirA 10 µM	38470 76925	Monomer, biotinyl-5′-AMP bound Dimer, biotinyl-5′-AMP bound	38466 76931
Apo-*Sa*BirA D200E 1.4 µM	37916	Monomer	37905
Holo-*Sa*BirA D200E 1.4 µM	38488	Monomer, biotinyl-5′-AMP bound	38479
Apo-*Sa*BirA D200E 11.25 µM	37919	Monomer	37905
Holo-*Sa*BirA D200E 11.25 µM	38484	Monomer, biotinyl-5′-AMP bound	38479
Apo-*Sa*BirA D200E 22.5 µM	37912	Monomer	37905
Holo-*Sa*BirA D200E 22.5 µM	38488	Monomer, biotinyl-5′-AMP bound	38479
Apo-*Sa*BirA D200E 45 µM	37915	Monomer	37905
Holo-*Sa*BirA D200E 45 µM	38489	Monomer, biotinyl-5′-AMP bound	38479
Apo-*Sa*BirA D200E 90 µM	37915	Monomer	37905
Holo-*Sa*BirA D200E 90 µM	38486	Monomer, biotinyl-5′-AMP bound	38479

Supplementary Table 2. Primers used in this study.

Primer	Sequence	Function	source
B481	5′ -GGTTGCTAATAATGAAGGTATAGAA GCAATAATATGTGG-3′	D200E mutagenesis	This study
B482	5′ -CCACATATTATTGCTTCTATACCTTCAT TATTAGCAACC -3′	D200E mutagenesis	This study
B200	5′-GGTATAGAAGCAATAATATGT GG-3′	Internal BPL sequencing	
Lambda P1	5'-GGCATCACGGC AATATAC-3'	attp-λ PCR screening primer	St. Pierre et al, 2013
Lambda P2	5'-ACTTAACGGCTGACATGG-3'	attp-λ PCR screening primer	St. Pierre et al, 2013
Lambda P3	5'-GGGAATTAATTCTTGAAGACG-3'	attp-λ PCR screening primer	St. Pierre et al, 2013
Lambda P4	5'-TCTGGTCTGGTAG CAATG-3'	attp-λ PCR screening primer	St. Pierre et al, 2013

Supplementary Table 3. Oligonucleotide sequences used for qRT-PCR to measure repression of bioD and bioY by *Sa*BPL.

Target gene	Primer name	Primer sequence 5'-3'	Source
S. aureus bioD	qSA2716_F	GCAAGGTGTGGTGATACAGG	Satiaputra et al, 2018
	qSA2716_R	ACACGTGGTCATCGAGTTTG	Satiaputra et al, 2018
S. aureus bioY	qSA2552_F	AATGGCAAGCCAGCAACTAC	Satiaputra et al, 2018
	qSA2552_R	GGATTGGTACCGGTAATTCCA	Satiaputra et al, 2018
S. aureus 16s rRNA	qSA0002_F	GAACCGCATGGTTCAAAAGT	Satiaputra et al, 2018
	qSA0002_R	CGTAGGAGTCTGGACCGTGT	Satiaputra et al, 2018

Supplementary Table 4. Oligonucleotide sequences employed for EMSA analysis of *Sa*BPL DNA binding activity.

Oligo name	**Sequence 5'-3'**	**Description**	**Source**
DS-SaBioO oligo 1	CCTTAAATGTAAACTTTTATAATT AATAAGTTTACATTTAAG	Top strand oligo containing *sabioO* wildtype sequence	Satiaputra et al, 2018
DS-SaBioO oligo 2	CCTTAAATGTAAACTTATTAATTA TAAAAGTTTACATTTAAGG	Bottom strand oligo containing *sabioO* wildtype sequence	Satiaputra et al, 2018
DS-SabioY oligo 1	AACTTATTGTAAACTTTTCATTTCT TAAAGTTTACAATGGTGCT	Top strand oligo containing *sabioY* wildtype sequence	Satiaputra et al, 2018
DS-SabioY oligo 2	AGCACCATTGTAAACTTTAAGAAA TGAAAAGTTTACAATAAGTT	Bottom strand oligo containing *sabioY* wildtype sequence	Satiaputra et al, 2018

Supplementary Table 5. Bacterial strains utilised within this study.

Strain	**Description**	**Function**	**Source**
E. coli DH5α	*sup*EΔ*lac*169 (p80*lac*ZΔM15) *hsdR17 rec*A1 *end* AA1 *gyr*A96 *thi*-1 *rel*A1	**Plasmid manipulation**	New England Biolabs, MA, USA
E. coli BL21-CodonPlus(DE3)-RIPL	*E. coli B F⁻ ompT DhsdS (rB⁻mB⁻) dcm⁺Tet^R gal λ(DE3) endA Hte [argU proL Cm^R] [argU ileY leuW Strep/Spec^R]*	**Protein expression strain**	Aglient Technologies, CA, USA
S. aureus NCTC 8325	Laboratory methicillin sensitive *S. aureus* with reference genome	**Parental *S. aureus* strain to *S. aureus* NCTC8325-B7**	American Tissue Culture Collection
S. aureus NCTC8325-B7	*S. aureus* NCTC8325 after 18 days passage BPL199 (*birA* D200E)	**Strain containing D200E mutation**	This study

Supplementary Table 6. Specific strains used for the *in vivo* beta-galactosidase reporter assay.

Strain	Integrations	Description	Source
JD26186 *birA::CAT*	*bioC::KanR birA::CAT*	JD28186 strain with N-terminal CAT cassette insertion (knockout)of its endogenous *birA*	Satiaputra et al, 2018
JD26186 *birA::CAT-SabioO-SaBPL*	*bioC::KanR birA::CAT (SabioO-lacZ)HK(placUV5-SaBPL)I*	JD26186 birA::CAT strain with SaBioO lacZ reporter chromosomally integrated at HK022 att site, and placUV5-SaBPL(wildtype) cassette chromosomally integrated at lambda att site.	Satiaputra et al, 2018
JD26186 *birA::CAT-SabioY-SaBPL*	*bioC::KanR birA::CAT (SabioY-lacZ)HK (placUV5-SaBPL)I*	JD26186 birA::CAT strain with SaBioY-lacZ reporter chromosomally integrated at HK022 att site and placUV5-SaBPL(wildtype) cassette chromosomally integrated at lambda att site.	Satiaputra et al, 2018
JD26186 *birA::CAT-SabioO-SaBPL D200E*	*bioC::KanR birA::CAT (SabioO-lacZ)HK (placUV5-SaBPL D200E)I*	JD26186 birA::CAT strain with SaBioO-lacZ reporter chromosomally integrated at HK022 att site, and plac-UV5-SaBPL (D200E) cassette chromosomally integrated at lambda att site.	This study
JD26186 *birA::CAT-SabioY-SaBPL D200E*	*bioC::KanR birA::CAT (SabioY-lacZ)HK (placUV5-SaBPL D200E)I*	JD26186 birA::CAT strain with SaBioY-lacZ reporter chromosomally integrated at HK022 att site, and plac-UV5-SaBPL (D200E) cassette chromosomally integrated at lambda att	This study

Supplementary Table 7. Plasmids utilised in this study.

Plasmid	Description	Source
pGEMT-*Sa*BPL(6xHis)	pGEMT plasmid containing *saBPL* with 6x his-tag	Pendini et al, 2008
pGEMT-*Sa*BPL-D200E-(6xHis)	pGEMT plasmid containing *saBPL* D200E with 6x his-tag	This study
pIT4_TL_152002	Chromosomal integration plasmid (λ-attP, Tc^R, R6Kγ ori, *ccdB,* pUC ori)	St. Pierre et al, 2013
pIT4_TL_SaBPL (D200E)	plac-UV5 fused with SaBPL (D200E) sequence cloned into pIT4_TL_152002	This study

References:

1. Pendini, N. R., Yap, M. Y., Polyak, S. W., Cowieson, N. P., Abell, A., Booker, G. W., Wallace, J. C., Wilce, J. A., and Wilce, M. C. J. (2013) Structural characterisation of *Staphylococcus aureus* biotin protein ligase and interaction partners: An antibiotic target. *Protein Sci* **22**, 762-773
2. Lee, K. J. (2019) Sulfonamide-based inhibitors of biotin protein ligase as new antibiotic leads.
3. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein engineering* **8**, 127-134
4. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. ACS Publications
5. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., and Söding, J. (2011) Fast, scalable generation of high‐quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**, 539
6. Okonechnikov, K., Golosova, O., Fursov, M., and team, U. (2012) Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166-1167