The structure and function of Biotin Protein Ligase:

A focus on *Staphylococcus aureus, Saccharomyces cerevisiae, Candida albicans* and *Homo sapiens.*

Nicole Renee` Pendini, B. Sc. Honours (University of Adelaide)

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

May, 2009

School of Molecular and Biomedical Sciences
Discipline of Biochemistry
University of Adelaide
South Australia

Index

Abbreviations

ACC: Acetyl CoA Carboxylase

AMP: Adenosine monophosphate

Apo: Unliganded enzyme

ATP: Adenosine triphosphate

BCCP: Biotin Carboxyl Carrier Protein

BirA: Biotin inducible repressor

BPL: Biotin Protein Ligase

Bt: Biotin

BtOH-AMP: Biotinol-5'-adenosine monophosphate

CaBPL: *Candida albicans* Biotin Protein Ligase

EcBPL: *Escherichia coli* Biotin Protein Ligase

HCS: Holocarboxylase synthetase

Holo: Ligand bound enzyme

MCD: Multiple carboxylase deficiency

MR: Molecular replacement

MRSA: Methicillin resistant *Staphylococcus aureus*

PC: Pyruvate carboxylase

PDB: Protein Data Bank

PDBID: Protein Data Bank identification code

PhBPL: *Pyrococcus horikoshi* Biotin Protein Ligase

R.M.S.D: Root mean square deviation

SaBPL: *Staphylococcus aureus* Biotin Protein Ligase

ScBPL: *Saccharomyces cerevisiae* Biotin Protein Ligase

SAXS: small angle X-ray scattering

VRSA: Vancomycin resistant *Staphylococcus aureus*

Abstract and summary of thesis for Nicole Renee Pendini

Biotin Protein Ligase (BPL) is an essential enzyme responsible for the covalent attachment of biotin to a specific lysine residue of biotin-dependent carboxylases, transcarboxylases and decarboxylases. Due to the fundamental processes that these enzymes are involved in such as lipogenesis, amino acid catabolism and gluconeogenesis, much research has been conducted on these enzymes. Studies encompassing structural, mutational and catalytic functions of these enzymes have lead to novel drug developments for the treatment of obesity, diabetes, metabolic syndrome, bacterial and fungal infections.

As BPL is required for activation of these enzymes by biotinylation, it is believed that it too could be targeted in a similar way to produce novel therapeutics. To date, the most characterised BPLs are from the Gram-negative bacteria *Escherichia coli* and the archea *Pyrococcus hirokoshii*. However minimal information is known about other forms of clinically important bacterial species or eukaryotic forms of this important enzyme.

Through my candidature I have compiled a thorough literature review summarised as chapter 1: Introduction. Furthering this literature analysis, a human BPL model was generated with aid of BPL structural co-ordinates already deposited in the protein data bank (PDB), thus allowing focus on human BPL mutations that cause multiple carboxylase deficiency (chapter 2). I have solved the structure of BPL from the clinically important pathogenic bacteria *Staphylococcus aureus.* This was performed in several ligand-bound and non-bound states (chapters 3 and 4). A novel high-throughput assay was

developed to test BPL activity. This assay allow testing of compounds that could potentially inhibit the BPL from *Candida albicans* (a species responsible for invasive fungal infections) (chapter 5). Large amounts of highly purified BPL from *Saccharomyces cerevisiae* allowed for the first structural analysis of a eukaryotic BPL (Chapter 6). The work has been summarised by a general discussion and future directions for the project (Chapter 7).

Thesis layout:

The thesis will be presented as a series of manuscripts either published or intention to be submitted for publication. Each manuscript will form a self-contained chapter with its own references. Included will be an overall general introduction and general discussion to link the information from these papers so as to present a uniform body of research conducted during candidature.

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

This work contains no material which has been accepted for the award of any other degree or diploma 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.

This thesis contains published work and/ or work prepared for publication of which some has been co-authored. In this case papers have joint and multiple authorship and therefore are accompanied by a statement of contribution (in terms of the conceptualisation and contribution of the work) by the candidate and other authors. Authors are required to sign and give permission for the paper to be included in the thesis.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

Nicole Renee` Pendini Date

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

Chapter 2:

Nicole R. Pendini, Lisa M. Bailey, Grant W. Booker, Matthew C. Wilce, John C. Wallace & Steven W. Polyak. (2008) Distant relatives of Biotin Protein Ligase aid in understanding multiple carboxylase deficiency. *Biochemica Biophysica Acta – Proteins and Proteomics*, Jul-Aug 2008, 1784(7-8), 973-82. PMID: 18442489

Chapter 3:

Pendini NR, Polyak SW, Booker GW, Wallace JC, Wilce MC (2008) Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from *Staphylococcus aureus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2008 Jun 1;64(Pt 6):520-3, 2008 May 23. PMID: 18540065

Chapter 4:

Nicole R. Pendini, Steven W. Polyak, Grant W. Booker, John C. Wallace, & Matthew C. Wilce, Crystal structures of apo and liganded Biotin Protein Ligase from *Staphylococcus aureus*: towards the development of new antibiotic for MRSA. To be submitted for publication.

Chapter 5:

Nicole R Pendini; Lisa M Bailey; Grant W Booker; Matthew C Wilce; John C Wallace; Steven William Polyak (2008) Biotin Protein Ligase from *Candida albicans:* Expression, purification and development of a novel assay, Arch Biochem Biophys. 2008 Nov 15;479(2):163-9. Epub 2008 Sep 11, PMID: 18809372

Chapter 6:
Nicole R. Pendini, Nathan Cowleson, John C. Wallace, Grant W. Booker, Matthew C. Wilce & Steven W. Polyak (2008) Characterisation of the domain structure of yeast biotin ligase and its complexes by small-angle X-ray scattering and molecular modelling. Submitted to JBC tracking number JBC/2009/013490.

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.

Communications and Presentations

Pendini, N. R., Cowieson, N., Wallace J.C., Booker G.W., Wilce, M.C and Polyak, S.W., The characterisation of the domain structure of yeast biotin ligase and its complexes by small-angle X-ray scattering and molecular modelling (2009), Biomolecular Dynamics and Interactions Symposia: Protein folding, function and assembly. Abs:P6.

Pendini, N. R., Polyak, S.W., Booker G.W., Wallace J.C. and Wilce, M.C., *Staphylococcus aureus* Biotin Protein Ligase as a novel antibiotic target (2009), 34th Lorne Conference on Protein Structure and Function. Abs#157.

Pendini, N. R., Polyak, S.W., Booker G.W., Wallace J.C. and Wilce, M.C., *Staphylococcus aureus* Biotin Protein Ligase as a novel antibiotic target (2008), University of Adelaide, School of Molecular and Biomedical Sciences Research Symposia. **Awarded** $200 for best poster presentation.

Pendini, N. R., University of Adelaide, School of Molecular and Biomedical Sciences Research Symposia, Scientific Image competition: **Awarded** $50 for best image.

Pendini, N. R., Polyak, S.W., Booker G.W., Wallace J.C. and Wilce, M.C., *Staphylococcus aureus* Biotin Protein Ligase as a novel antibiotic target (2008) Melbourne Protein Group meeting, Bio21, Victoria, Australia **Awarded** $50 for best poster presentation.

Pendini, N. R., Polyak, S.W., Booker G.W., Wallace J.C. and Wilce, M.C., Biotin Protein Ligase as a novel antibiotic target (2008), Monash University, Department of Biochemistry and Molecular Biology Annual Postgraduate Research Conference, invited speaker. **Awarded** $200 for best presentation. Abstract S04.

Pendini, N. R., Polyak, S.W., Booker G.W., Wallace J.C. and Wilce, M.C., Biotin Protein ligase as a novel antibacterial target: a focus of *Staphylococcus aureus* (2008) Sicily, Italy, 40th Course for the International School of Crystallography: From Molecules to Medicine: Integrating Crystallography in Drug Discovery, Pos 67. **Awarded** student travel scholarship € 600.00.

Pendini, N. R., Polyak, S.W., Booker G.W., Wallace J.C. and Wilce, M.C., Biotin Protein Ligase as a novel antibacterial target (2008), 33rd Lorne Conference on Protein Structure and Function. Abs#361. **Awarded** student travel award $100.

Pendini, N. R., Polyak, S.W., Bailey, L.M., Booker G.W., Wilce, M.C. and Wallace J.C., Purification and characterisation of Biotin Protein Ligase from *Candida albicans* (2007), Monash University, Department of Biochemistry and Molecular Biology Annual Postgraduate Research Conference.

Pendini, N. R., Polyak, S.W., Bailey, L.M., Booker G.W., Wilce, M.C. and Wallace J.C., Purification and characterisation of Biotin Protein Ligase from *Candida albican* (2007), 32rd Lorne Conference on protein structure and function Abs#238. **Awarded** student travel award $100.

Pendini, N. R., Polyak, S.W., Swift, R., Booker G.W., Wilce, M.C. and Wallace J.C., Probing the importance of the N-terminal region of Human Biotin Protein Ligase (2007), The University of Adelaide Molecular and Biomedical Sciences School Symposia: **Finalist** in the scientific image competition for "BirA crystal images".

Pendini, N. R., Polyak, S.W., Booker G.W., Wilce, M.C. and Wallace J.C., Determination of the Structure of Yeast Biotin Protein Ligase: Implications as a Novel Antifungal Drug Target (2006), Australian Society for Medical Research SA division Scientific Meeting, invited presenter. **Awarded:** $400 for Best oral presentation by an ASMR student member.

Pendini, N. R., Polyak, S.W., Booker G.W. and Wallace J.C., Probing the importance of the N-terminal region of Human Biotin Protein Ligase (2006), 31st Lorne Conference on Protein Structure and Function. **Awarded** student travel award $100.

Pendini, N. R., Polyak, S.W., and Wallace J.C., 2005-Combio
The purification of Eukaryotic Biotin Protein Ligase.

Pendini, N. R., Polyak, S.W., and Wallace J.C., The purification of yeast Biotin Protein Ligase (2005). The East Coast Protein Structure Conference.

Pendini, N. R., Polyak, S.W., and Wallace J.C., (2005), The purification of yeast biotin protein ligase, Australian Society for Medical Research SA division.

Pendini, N. R., Polyak, S.W., Swift, R., and Wallace J.C., A survey of human Biotin domains as a substrate for Biotin Protein Ligase (2005) 30th Lorne Conference on Protein Structure and Function. **Awarded** student travel award $100.

Co-author poster presentations:

Ng, B., Pendini N.R., Tieu W., Kuan K., Morona R., Abell A., Wilce M.C.J., Wallace J.C., Polyak S.W. and Booker G.W. (2008) Discovery of biotin protein ligase as a novel class of antibiotic. University of Adelaide, School of Molecular and Biomedical Sciences Research Symposia, Pos 12

Invited Presentations at National Conferences
Polyak, S.W., Tieu, W., Ng, B., Pendini, N.R., Kuan, K., Morona, R., Booker, G.W., Wilce, M.C., Wallace, J.C., Abell, A.D. (2008) Inhibitors of biotin protein ligase: A novel class of antibiotics for the treatment of *Staphylococcus aureus* Proc. Aust. Health and Medical Research Congress Abstract 153

Ng, B., Pendini N.R., Tieu W., Kuan K., Morona R., Abell A., Wilce M.C.J., Wallace J.C., Polyak S.W. and Booker G.W. (2008) Discovery of biotin protein ligase as a novel class of antibiotic. ComBio2008, Pos WED-024

Ng, B., Pendini, N., Tieu, W., Kuan, K., Morona, R., Wallace, J. C., Wilce, M., Abell, A., Booker, G. W. & Polyak, S. W. (2008) Inhibitor of biotin protein ligase: a novel class of antibiotics for the treatment of *Staphylococcus aureus*. Australian Society for Medical Research SA division Scientific Meeting

Mayende, L., Swift, R. D., Bailey, L. M., Pendini, N. R., Wallace, J. C. & Polyak, S. W. (2008) Domain structure of human holocarboxylase synthetase: evidence of an interaction between the N-terminal and C-terminal halves Proc. Lorne Conference on Protein Structure and Function, Pos 255

Polyak, S. W., Stojkoski, C., Pendini, N. R., Booker, G. W. & Wallace, J. C. (2007) Biotin protein ligase: A novel antibiotic target Proc. Lorne Conference on Protein Structure and Function, Pos 323

Acknowledgements

I would like to thank Professor John Wallace for taking a chance on my risky ambition to study X-ray crystallography as a PhD project and Dr Steven Polyak for all his advice and lab training that have allowed me to grow from a standard graduate student into a successful scientist. To Associate Professor Matthew Wilce, for giving me the opportunity not only to study crystallography but a range of techniques that would not have been possible without your help. To Dr Nathan Cowieson for the SAXS data collection and analysis and to Dr Grant Booker and Dr Andrew Abell for their support for the BPL project.

To all the past and present members of team BPL, namely Lisa Bailey, Lungisa Mayende, Belinda Ng, Daniel Bird, Ruby Ivanov, Rachel Swift and Fiona Whelan as well as other supportive members of the Wallace lab including Briony Forbes, Carlie Delaine, Claire Alvino and Kerrie McNeil. To past and present members of the Wilce lab including Jackie Wilce, Marlies Loescher, Sumay Ng, Jason Schmidberger, Corrine Porter, Julian Vivian, Min Yin Yap, Andrew Sivakumaran, Henry Kim, Simone Beckham and Edward Cummings. All these people mentioned have contributed extremely useful advice during my candidature and I really have appreciated all that I have learned from each and every one of you. Thanks also goes to the Booker, Rossjohn and Whisstock labs for being generous and allowing me to borrow equipments and reagents in order to complete my experiments in record time!

Finally, thanks to my family for being so supportive from afar and always giving me a place that I can call home. Thank you so much, miss you, love you.

I wish to dedicate this work to Diana Visentin and Lucy Nunn,

my reasons for studying so hard in a field that can be so thankless.

Chapter 1

General Introduction

The overall objective of this multidisciplinary, collaborative research project is to discover new compounds that can be used for the treatment of infection by pathogenic bacteria and fungi. To achieve this our research team is focussed upon the novel drug target, Biotin Protein Ligase (BPL). The aim of the studies presented in this thesis is to further our understanding of the structure and function of this important metabolic enzyme using structural biology, biochemistry and *in silico* drug screening. This information is an important aspect to our drug-discovery programme. Of note, some literature has been omitted from the introduction to avoid repetition with the review presented as Chapter 2.

Antimicrobial chemotherapy of pathogenic bacteria

There are many reports of antibiotic resistance and how we have moved into the "super-bug" era, yet little investigation is being performed to generate new drugs for these deadly organisms. Most antibiotics on the market target one of 4 major cellular processes; cell wall and membrane synthesis, protein biosynthesis, DNA replication and folate co-enzyme biosynthesis (1). The majority of "new" products appearing on the market are simply variants of existing treatments that exert their effect on the same target making the acquisition of resistance less difficult for the bacteria. Despite the need for new compounds with novel modes of action, large pharmaceutical companies are not embracing this challenge. Resistance has been observed to every class of antibiotic, regardless of whether it was derived from natural or synthetic origins (1). The first example of this was observed with penicillin resistant *Staphylococcus aureus* in the 1940s. The second line of defence against bacterial infection was β-lactam where resistance developed within 10 years. The development of methicillin was thought to deliver the 'magic bullet' but in 1986, Methicillin-resistant *S. aureus* (MRSA) was observed. This was followed with vancomycin resistance (2) in 1996.

In recent years there have been several first in class antibiotics approved for use in the clinic. In 2001, linezolid (Zyvox®, Pfizer), that targets protein synthesis, became the first new antibiotic to receive FDA approval in over 30 years. This was followed by daptomycin (Cubicin®, Cubist) in 2003 and tygacycline (Tygacil®, Wyeth) in 2005 (3). Whilst several new products are in the pipeline, it is well recognised that more effort is required to stay ahead of drug resistance (1). One important strategy to combat drug resistance is to deliver new agents that combat new drug targets where there is no pre-existing resistance.

Between 1995-2001 GlaxoSmithKline investigated over 300 genes in *Staphylococcus aureus* as potential targets for antibiotics. Through genetic studies they showed that 160 were essential for survival. 70 of these genes were then further tested in high-throughput screening (4). Two of the targets identified were Biotin Protein Ligase (BPL) and Acetyl CoA carboxylase (ACC). These candidates were screened against random compound libraries (260,000-530,000 compounds) to identify potential drug leads. However no hits were identified to these targets (4), suggesting a more directed approach is required. In this thesis I will present our efforts towards new BPL inhibitors using X-ray crystallographic structures with the aim of structure-based drug design.

Pathogenic yeast, fungi and moulds

Yeast, fungi and moulds possess hardy cell walls that form a protective barrier to toxins and, consequently, antifungal compounds (5). Therefore developing new agents for the treatment of pathogens belonging to this phylum is considered more challenging than bacteria. Furthermore treating invasive fungal infections, most commonly from *Candida spp.* and *Aspergillus spp*., is becoming increasingly difficult due to the prevalence of new pathogenic species, slow diagnosis, variable drug bioavailability, toxicity of current drugs and development of resistance. Amphotericin B deoxycholate, which associates with the essential cell membrane sterol, ergosterol leading to K^+ leakage and cell death, has been used for the treatment of fungal infections (6). However concerns over adverse side-effects such as fever, chills, nausea and vomiting, have been reported (7). In the 1990's two azole based compounds were released into the clinic; fluconazole and intraconazole but

problems with narrow antifungal activity and poor absorption have limited their use (8). More recently second generation triazoles have been developed that inhibit the cytochrome P450 dependent conversion of lanosterol to ergosterol (9). This leads to the accumulation of the toxic sterol precursor 14α-demethylsterol and subsequent destruction of the fungal cell membrane. Voriconazole was the first triazole to be approved in 2002 (10) followed by posaconazole in 2006 (11) while ravuconazole is still in clinical trials (12). These second generation triazoles are generally well tolerated though there have been several reported side effects including visual disturbances, fever, chills, headache, vomiting and abnormal liver function. They also possess a broader-spectrum of activity over the first generation azoles, and thereby are useful for the treatment of opportunistic fungi, including some fluconazole-resistant strains (13).

In addition, in 2001 the new class of antifungals known as echinocandins were introduced. This functions by inhibition of (1,3)-β-D-glucan synthetase required for production of the main fungal cell wall polysaccharide constituent beta-glucan. Currently there are three forms available for intra-venous use including caspofungin, micafungin and anidulafungin. These are broadly active against *Candida* spp. including strains resistant to azoles and polyenes. However, side effects similar to those presented for the second-generation azoles are observed (14). Of concern are resistant isolates that have already been identified with mutations in the target enzyme, making them resistant to inhibitor binding (15). Though the second-generation azoles and echinocandins are still important in the clinic, invasive isolates are being identified with adaptive resistance to these new drugs (16).

Combating drug resistance

One method of controlling pathogens is to use drugs that target essential pathways required for survival. Through a detailed understanding of the important proteins in these pathways, new drugs that work through previously unexplored targets could potentially be identified. Such agents are needed to combat drug-resistance, as microbes are unlikely to have pre-existing resistance mechanisms, thus making aversion to treatment more difficult. Throughout this thesis, I will focus my investigations on biotin protein ligase (BPL) as an exciting new drug target. Here I have reviewed the literature describing BPL and highlight its critical role in biotin metabolism in bacteria and yeast (the role of BPL in mammals is covered in depth in chapter 2).

Biotin

Biotin, otherwise known as vitamin H or B_7, is an important micronutrient in all organisms. Biotin cannot be synthesised by mammals and some micro-organisms but is absorbed in the body from one's diet and from intestinal bacteria that can produce biotin *de novo* (17). Biotin is actively absorbed into the circulation and enters mammalian cells via sodium-dependent multivitamin transporters. The most extensively characterised and understood function for biotin in cells is to serve as a cofactor for biotin-dependent enzymes. There are three classes of reactions in which biotin-dependent enzymes can be sub-divided. 1) For the carboxylases, biotin becomes carboxylated in a reaction requiring ATP, Mg^{+2} and bicarbonate. Here the carboxylate anion is transferred onto acceptor substrates such as pyruvate, propionyl-CoA, acetyl-CoA and β-methyl-crotonyl-CoA (18). 2) The decarboxylases mediate sodium transport in anaerobes and couple decarboxylation of oxalacetate, methylmalonyl-CoA and glutaconyl-CoA.

3) The transcarboxylases couple two reactions, for example the decarboxylation of oxalacetate to pyruvate followed by carboxylation of propionyl-CoA to methymalonyl-CoA (19).

The biotinylation reaction mechanism

Biotin protein ligase (BPL) is the enzyme responsible for specifically attaching the biotin onto the Biotin Carboxyl-Carrier Protein (BCCP). This reaction is related to that of amino acyl-tRNA synthetases (20,21) and lipoyl ligase (22,23) as the reaction proceeds through the formation of an adenylated intermediate (24). Biotin is attached to the biotin domain of biotin-dependent carboxylases in a two-step reaction whereby energy in the form of ATP is used to activate biotin in the presence of Mg^{+2}. The now activated biotin, a mixed anhydride biotinyl-5'-AMP, is attached to a specific lysine residue of an unbiotinylated (apo) domain of a biotin acceptor domain. It appears that the reaction mechanism and the protein: protein interaction between BPLs and biotin enzymes is conserved throughout biology, as it has been shown that BPLs are interchangeable between organisms (24-29) albeit poorly in some cases. This is possible due to the structural conservation between both the modifying enzymes and biotin accepting domains. BCCP homology at the primary structure level can be seen not only between biotin requiring enzymes across various species but also with non-biotin requiring enzymes such as the lipoyl enzymes (22). Computational analysis based on PSI-BLAST profiles, as well as secondary structure predictions of lipoylating and biotinylating enzymes, indicate that these modifying enzymes are evolutionarily related protein families containing a homologous catalytic module and therefore may share a common ancestral relationship through the use of an adenylated intermediate (22).

Biotin Protein Ligase

The attachment of biotin onto biotin-dependent enzymes is catalysed by the ubiquitous enzyme BPL, also known as the biotin inducible repressor, BirA, in *E. coli,* and as holocarboxylase synthase (HCS) in mammals. It was once believed a separate HCS existed for each of the carboxylases. However, with the availability of modern recombinant DNA technology and complete genome sequences, there is good evidence that only one biotin protein ligase gene is present in most bacteria, yeast and mammals, though there are significant differences between phyla. *Arabidopsis thaliana* and some other plants are exceptions that possess two HCS, one genes encoding a cytoplasmic isozyme and the other a chloroplast-targeted isozyme (30). Other species including *C. acetobutylicum, Lactococcus lactis, Halobacterium sp., Pyrococcus abyssii and Pyrococcus furiosus* also contain two genes for BPLs which differ by the inclusion or exclusion of an N-terminal regulatory region (31). Of all the BPLs, *E. coli* (EcBPL) is by far the most characterised and understood family member (24). A recent ensemble of BPL structures from the thermophilic archea *Pirococcus horikoshii* OT3 (43) have also provided new insights into the catalytic mechanism of BPLs (45).

Biotin metabolism in *E. coli* (BirA)

EcBPL is a truly bifunctional molecule that participates in both biotin attachment onto BCCP and also functions as a transcriptional repressor of the biotin biosynthetic operon (32). Thus this enzyme is commonly referred to as the Biotin inducible repressor (BirA). By combining these two activities into one single protein bacteria have developed exquisite regulation of biotin biosynthesis to meet cellular demand. Through structural biology, genetics and

biochemical analysis a detailed understanding of these processes has been obtained.

The structure of BirA

BirA (35.5 kDa) contains three distinct domains that have been determined at 2.3 Å resolution in 1992 through X-ray crystallography (PDBID: 1BIA). The monomeric structure measures 75 Å x 35 Å x 30 Å for the unliganded "apo" structure. The N-terminal 22-46 residues adopt a helix-turn-helix motif, a structure associated with DNA binding proteins (33). The central domain consists of five α-helices, 7 strands of mixed β-sheets as well as four poorly-defined loops that appear in pairs in the 3D structure. These loops consist of residues 110-128, 212-233 and 140-146 and 193-199 (figure 1a). The C-terminus consists of 6 strands which form a β-sandwich that seals the end of the enzyme and has been found to function in the transfer of biotin onto BCCP (33-35).

The crystal structure of BirA in complex with biotin was resolved to 2.4 Å resolution in 2001 (35). In the presence of biotin the loop consisting of residues 110-128 become ordered and encompassed the biotin cofactor (35) (figure 1b). Mutational analysis of this region showed reduced enzyme activity (36), highlighting the loop's importance. Kinetic experiments with these mutant enzymes showed a similar K_M for ATP, suggesting this region does not have a role in ATP binding though the region contained a sequence related to the nucleotide binding sequence GXGXXG. Structural studies have since revealed that the highly conserved GRGRXG sequence functions as a cap enclosing the biotin binding pocket as well as the ATP binding pocket. This is expected as the ATP and biotin moieties must be in close proximity to allow the

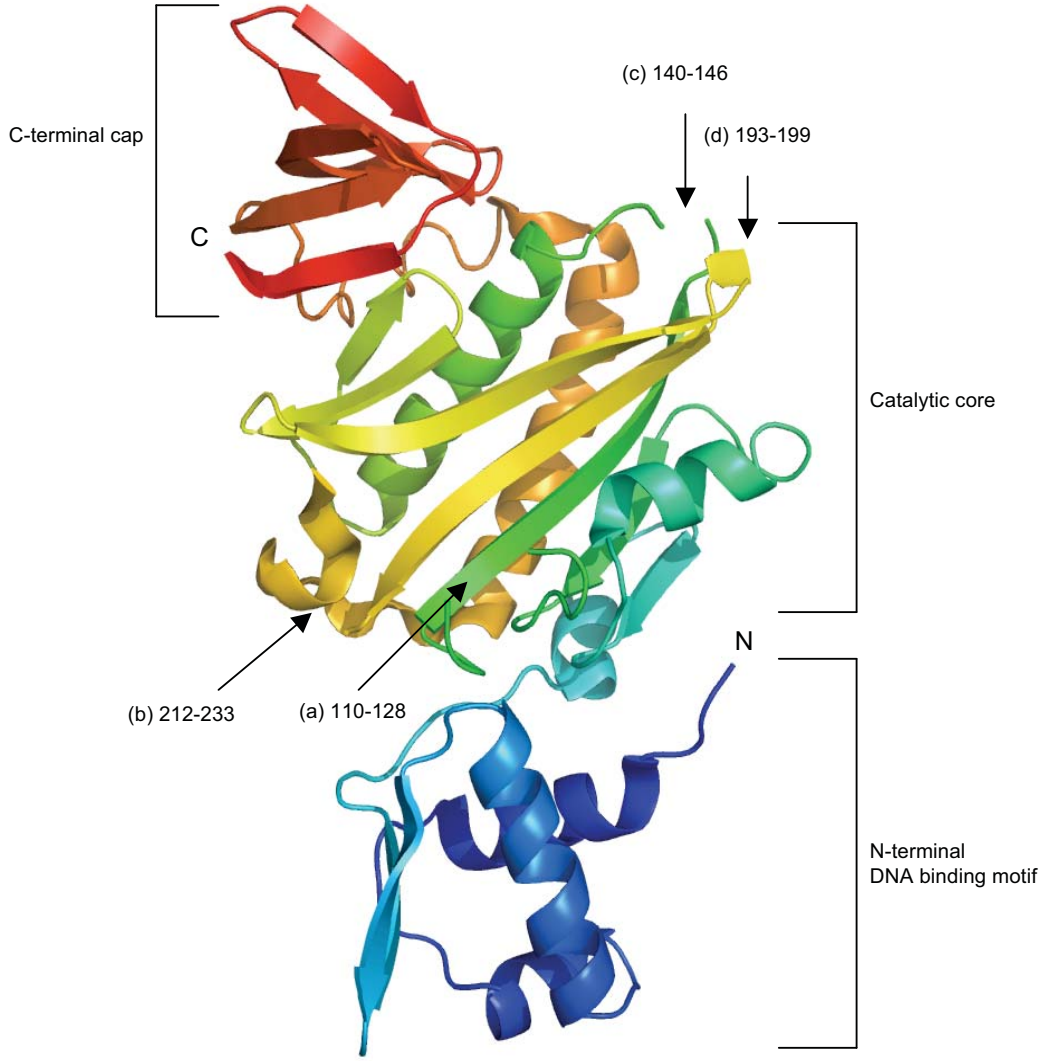


Figure 1a: Apo EcBPL (PDBID 1BIA)

The first structure of a BPL was that of *E. coli* solved in unliganded form by Wilson *et al* in 1992 by X-ray crystallography. The structure was determined to have 3 structural domains: 1) an N-terminal helix-turn-helix, known as a DNA binding motif (blue). 2) Central domain which has been shown through mutagenesis, to bind biotin and ATP and was predicted to be the catalytic core (green and yellow) and 3) C-terminal cap, shown through truncation studies to have some role in catalysis (red). In the apo structure, several surface loops were undefined in the electron density and presumed to be mobile (indicated by arrows with residue range) these loops appear in pairs a-b and c-d.

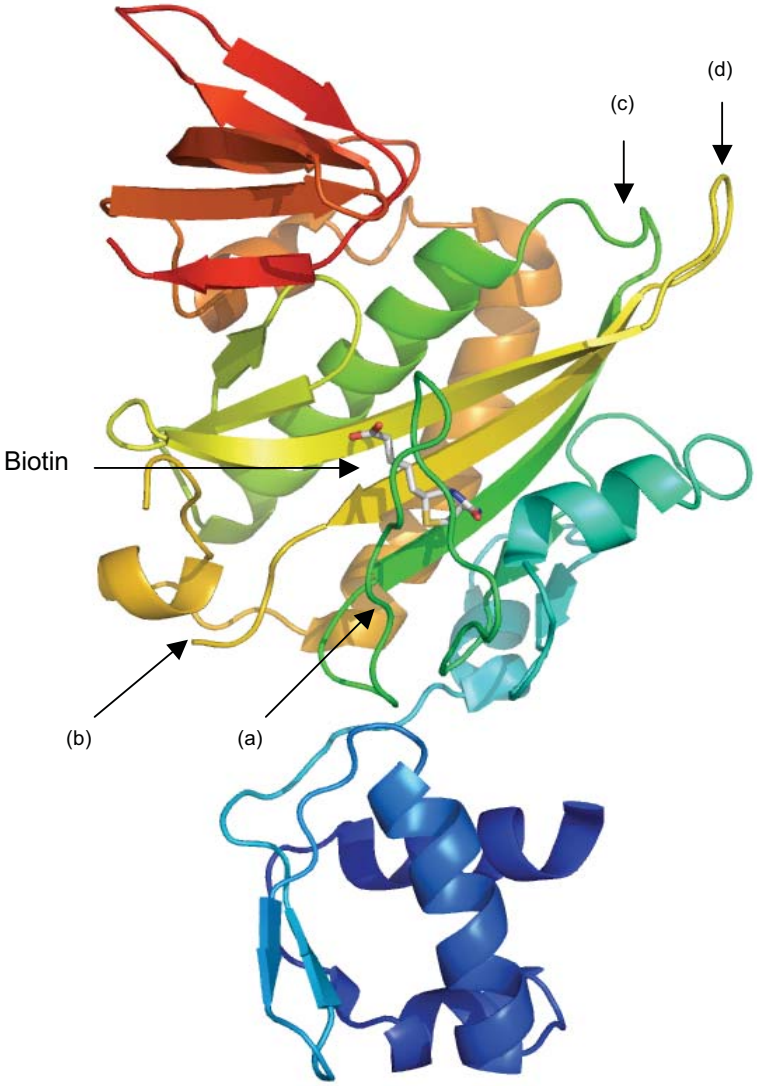


Figure 1b: Biotin bound EcBPL (PDBID 1HXD)

Almost a decade after the determination of the first structure of BirA came the 1st ligand bound structure of a BPL, again from *E. coli* co-crystallised with biotin by Weaver *et al* in 2001 (35). The structure showed that the surface loops that were disordered in the apo-structure (figure 2a) became ordered upon biotin addition. Biotin was shown to induce dimerisation, EcBPL loops (c) and (d) form part of the dimerisation interface. Loop (a) was shown to be involved in ligand binding (green). Loop (b) is still disordered.

formation of biotinyl-5'-AMP. Additionally mutations that affect enzymatic activity also map close to known biotin-contacting residues (33). In 2006, the structure of *E. coli* BPL co-complexed with the non-hydrolysable analogue of biotinyl-5'-AMP, biotinol-5'-AMP (37), was solved (figure 1c). This allowed the position of the nucleotide-binding site to be identified adjacent to the biotin-binding pocket (38) (figure 1d).

The helix-turn-helix N-terminal domain of BirA is not directly involved in the catalytic mechanism, as experiments on truncated BirA variants demonstrated synthesis of biotinyl-5'-AMP and transfer of biotin onto BCCP. The N-terminal domain of BirA is however essential for DNA binding and regulation of the biotin biosynthetic operon. In *E. coli*, the bio(ABFCD) locus contains five of the six genes required for biotin biosynthesis (39). The sixth gene (bioH) and the *birA* genes are located on separate regions of the *E. coli* chromosome. Synthesis of biotin is repressed by exogenous biotin concentrations greater than 40 nM (40). It is believed that two BirA monomers in complex with biotinyl-5'-AMP bind cooperatively with the biotin operator (41) although biotin has been shown, to a lesser extent, to activate dimerisation and DNA binding (42). Two holoBirA monomers bind 12-bp at the ends of a pseudo-palendromic 40-bp operator sequence. It is believed that the DNA bends slightly and BirA-biotin dimer undergoes a conformational change in order to accommodate specific interactions with one another, though this has yet to be determined structurally. As apo-BCCP accumulates in the cell it is able to compete with BirA dimerisation thereby inducing transcriptional derepression and utilising the bound biotinyl-5'-AMP for biotinylation. Further information describing BirA is found in Chapter 2.

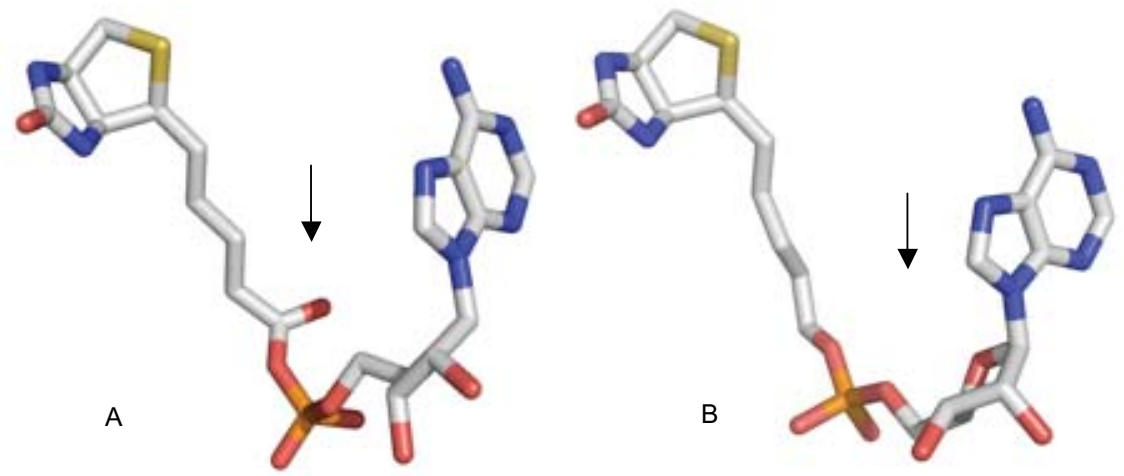


Figure 1c: Structures of biotinyl-5' -AMP vs biotinol-5'-AMP

(A) Illustrates the structure of biotinyl-5'-AMP in the conformation observed in the PhBPL structure (PDBID 2WQW) and (B) is the structure of biotinol-5'-AMP in the conformation observed in EcBPL (PDBID 2EWN). This is a synthetic derivative of biotinol produced by the formation of a phosphate-ether linkage with 5'-AMP which was shown to be non-hydrolysable by fluorescence spectroscopy and mass spectroscopy by Brown *et al* 2004 (37). The two structures are very similar with the only notable difference being the lack of the carbonyl group on biotin (arrows) therefore nucleophilic attack can not occur, and the second step of the biotin transfer reaction is prevented.

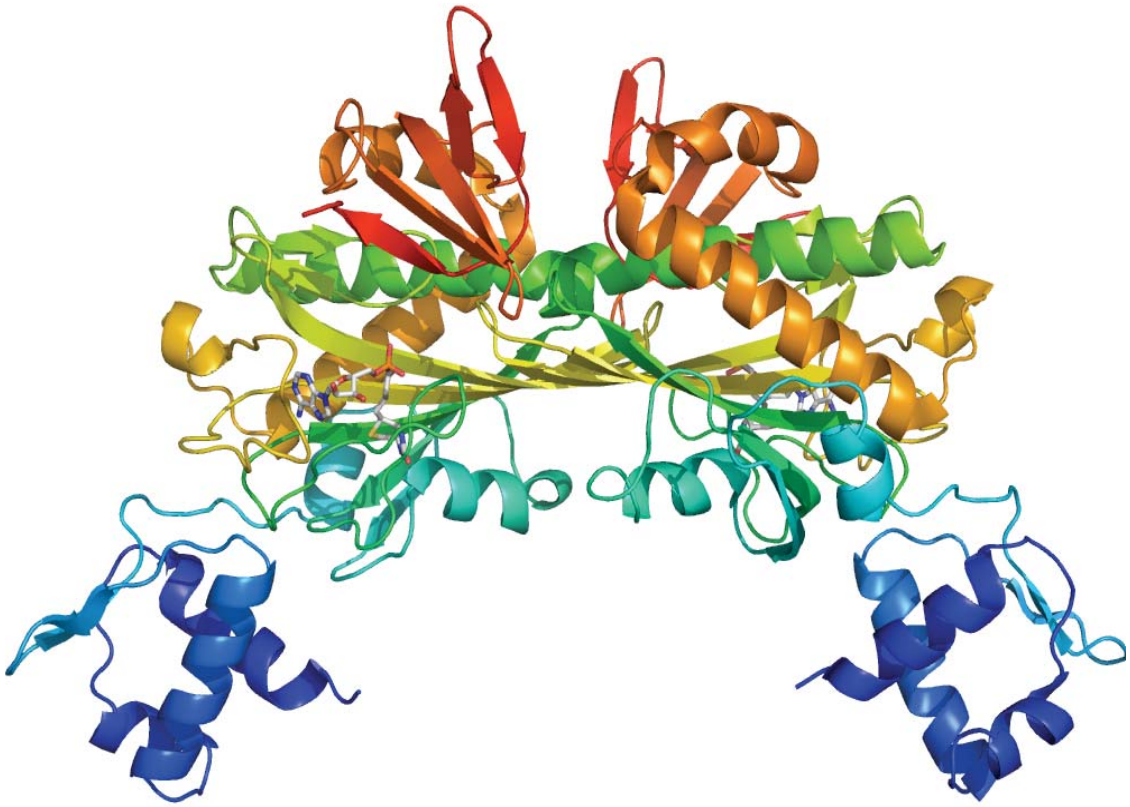


Figure 1d: Biotinol-5'-AMP bound EcBPL (PBD 2EWN)

The structure of EcBPL with the co-repressor analogue biotinol-5'-AMP (displayed in stick representation) shows in detail the dimerisation interface and complete biotin biosynthetic operon repressor. Loop (b) was shown to interact with the adenylate of biotinol-5'-AMP explaining why this was still disordered in the biotin bound structure (figure 2b). This dimer was also shown to be a tighter dimer than that induced by biotin alone by Wood *et al* (2006) (38).

The structure of BPL from *P. horikoshii* OT3 (PhBPL)

Until recently, the only structural information regarding BPLs came from *E. coli*. However, the structure of BPL from *Pyrococcus horikoshii* OT3 (an Archaeal hyperthermophile) has been resolved in several forms, including apo (1.6 Å, PDBID 1WQ7) or in complex with biotin (1.6 Å, PBDID 1WPY), ADP (1.6 Å, PDBID: 1WN1) and biotinyl-5'-AMP (1.45 Å, PDBID 1WQW). Unlike EcBPL, PhBPL is a constitutive dimer that is structurally distinct from BirA as both molecules employ different interfaces for dimerisation (43). Furthermore, PhBPL does not contain a DNA binding motif at the N-terminus (43) (figure 2a) suggesting a lack of the repressor activity associated with BirA. *Ph*BPL shares 31% sequence homology with the C-terminus of BirA and the structures show a R.M.S.D of 2.41 Å when C^{α} atoms are superimposed (43) (figure 2b). The ADP binding site was found to be adjacent to the biotin-binding pocket and held by electrostatic and polar interactions. Glycine residues at the bottom of the biotin pocket are essential for ligand binding as are the hydrophilic residues surrounding the inner pocket. The hydrophobic tail of biotin is further stabilised by interactions with hydrophobic residues at the entry site (figure 2c). Like EcBPL, several loops in the structure are poorly ordered in the uncomplexed enzyme but become more ordered upon ligand binding. This includes the disordered loop between β_2-β_3 that becomes structured upon ligand binding (43). This observation can also be observed in EcBPL where residues Arg116 and Arg118 undergo a conformational change that 'holds' biotin in the binding pocket (independent observation, N. Pendini chapter 4). Though analogies can be made between the structures solved thus far little more is known about the biology of *Ph*BPL in *P. horikoshii* OT3. Further information on PhBPL can be found in chapter 2.

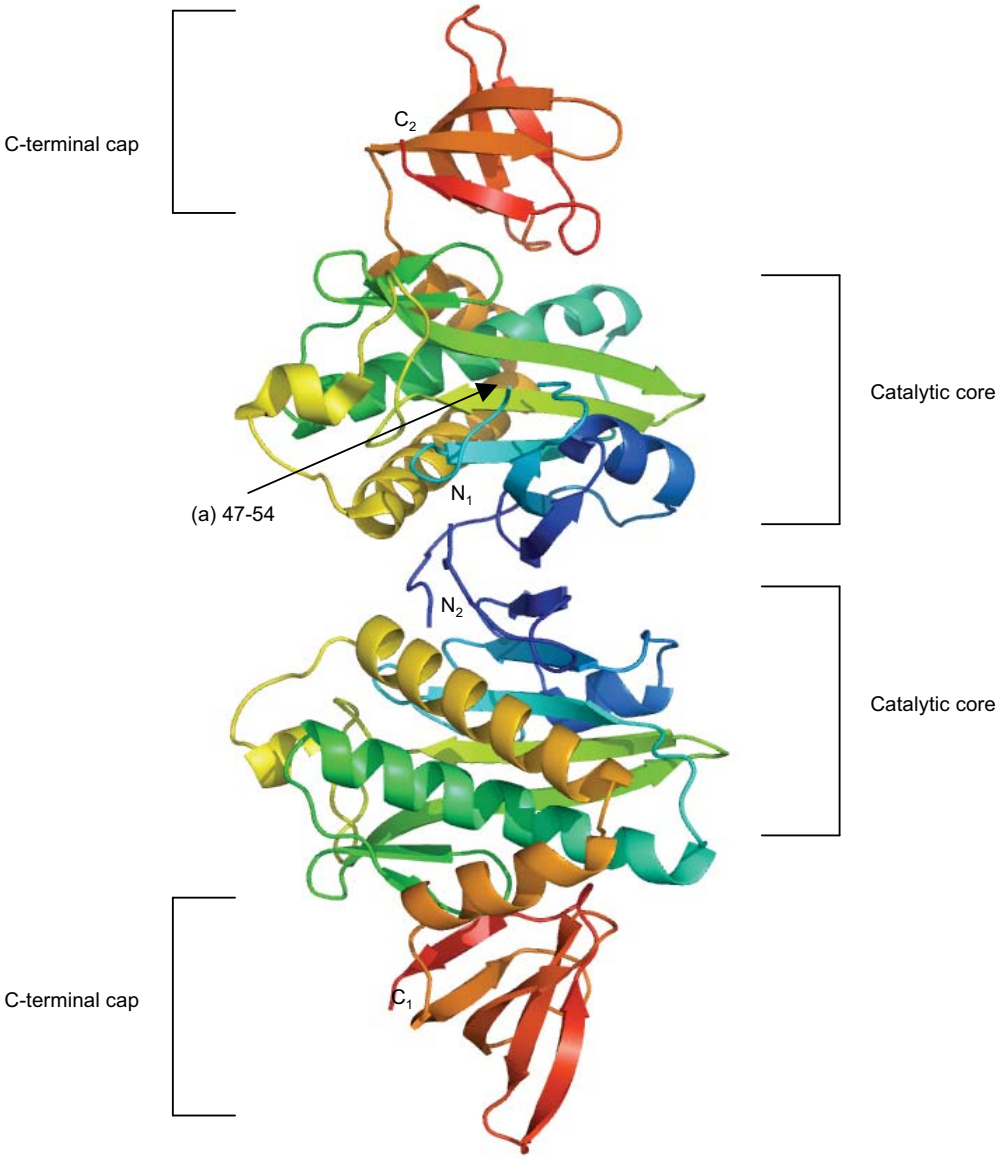


Figure 2a: Apo PhBPL (PDB 1WQ7)

Unliganded BPL from *Pyrococcus horikoshii* forms a constitutive homodimer. The dimer interface differs from that of EcBPL where the catalytic core forms the interface, whereas for PhBPL the interface is between the N-termini. PhBPL lacks a DNA-binding motif allowing this dimerisation event to be possible. Analogous to the apo EcBPL structure, loop (a) 110-124 (47-54 PhBPL numbering) was found to be unstructured by Bagautdinov *et al* 2005 (43).

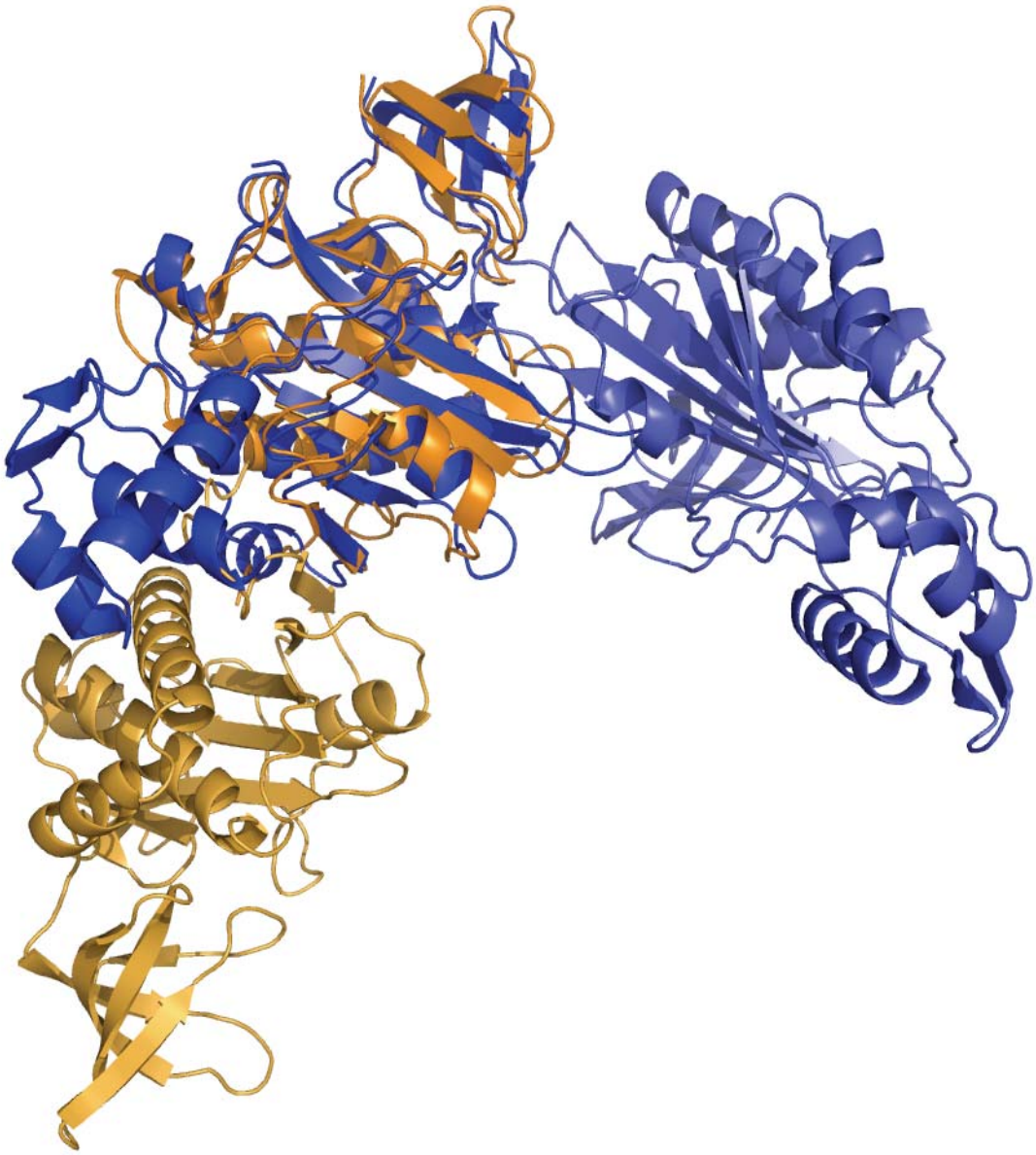


Figure 2b: Superposition of biotin bound EcBPL and PhBPL

The dimers of EcBPL (blue) and PhBPL (orange) superposed. When PhBPL is bound to biotin, the loop covering the active site becomes ordered and the catalytic core and C-terminal cap resemble EcBPL.

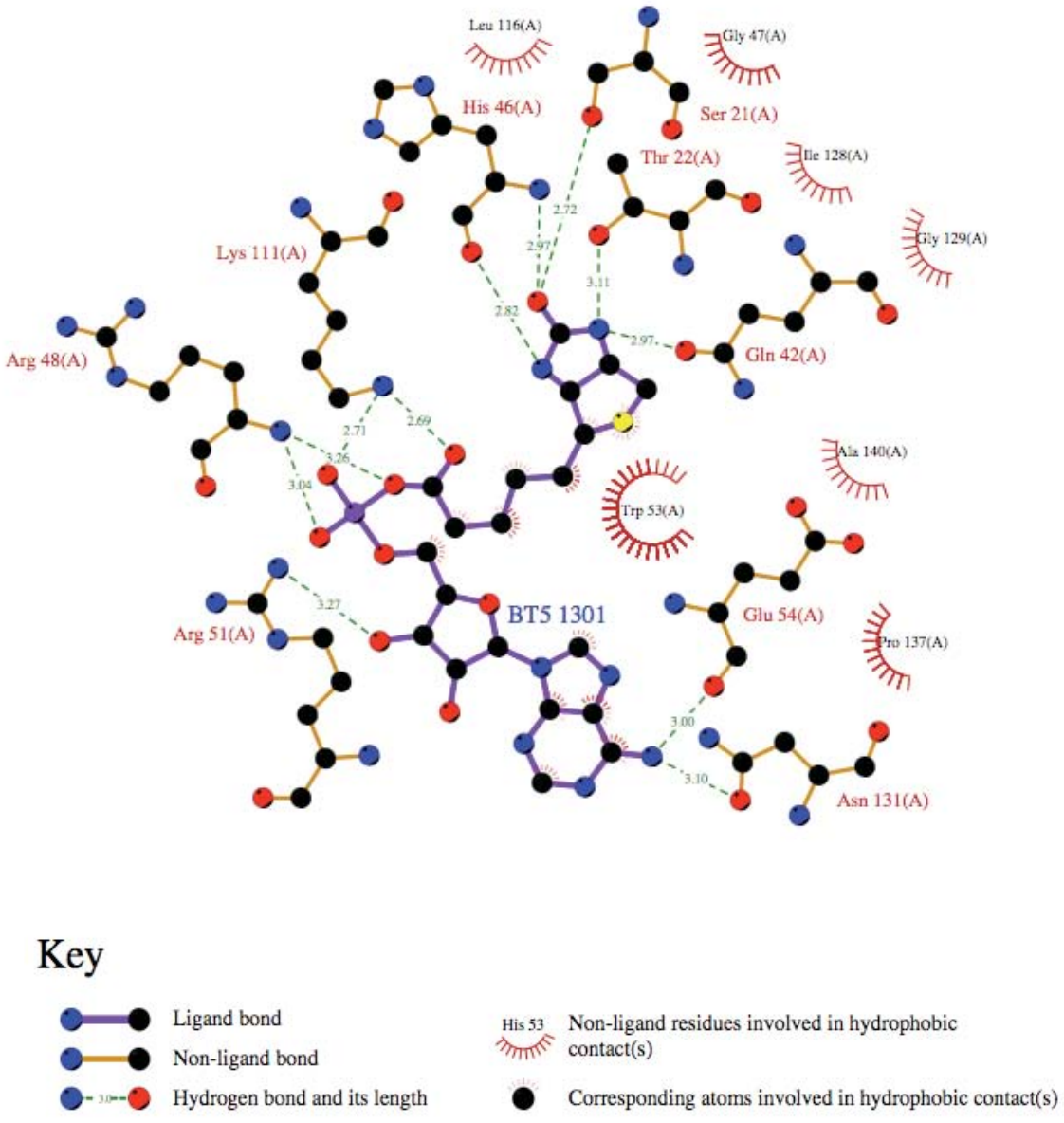


Figure 2c: Hydrogen bonds and hydrophobic interactions between PhBPL and biotinyl-5'-AMP (PDBID 1WQW)

The structure of PhBPL bound to biotinyl-5'-AMP was the first structure solved with the natural ligand. This provided important structural details of the important residues involved in bonding and interacting with the ligand and explained why the C-terminal cap was essential for biotinylation.

*Ph*BPL in complex with BCCP

Recently the structure of PhBPL in complex with BCCP was reported (44). This is the first structure of such a complex. To capture the complex by crystallography, mutants of the BPL were generated, namely R48A (PDBID 2EJG) and R48A with K111A (PDBID 2EJF). These residues were targeted as both were shown to be essential for hydrogen bonding to biotinyl-5'-AMP (PDBID 1WQW) (figure 2c). The C-terminal 73 residues from the biotin-containing gamma subunit of methymalonyl-CoA decarboxylase served as the biotin domain (45). The resulting structure showed that the binding interface between PhBPL and BCCP is analogous to the dimerisation interface in BirA (Figure 3). These structures have provided information as to important residues within the BPL structure that are involved in substrate recognition. As mentioned earlier, PhBPL does not contain an N-terminal DNA-binding motif required to regulate biotin biosynthesis. Whilst a mechanism addressing competition between substrate binding and homodimerisation in *E. coli* can be proposed, one must be aware that the biology between the two species appears to be very different. These differences are even greater when we compare what is known of bacterial and archael BPLs to that which is known of eukaryotic BPLs.

Biotin metabolism in yeast

Saccharomyces cerevisiae has served as a model eukaryotic organism for many years. Whilst there are some similarities in biotin metabolism between yeast and mammals there are also some important differences. *S. cerevisiae* is unable to synthesise biotin *de novo* and must acquire the vitamin from exogenous sources. Biotin is imported into the cell through high-affinity biotin transporters (Vht1p), which is a protein-coupled symporter located in the

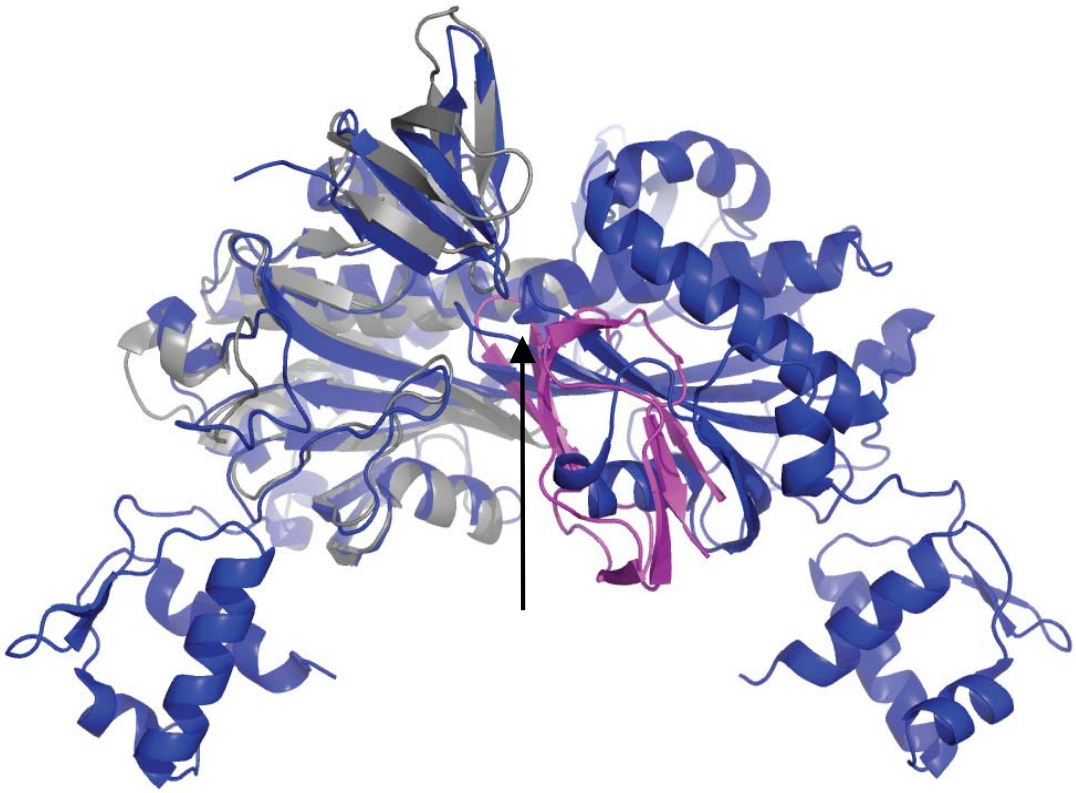


Figure 3: PhBPL:BCCP complex (PDBID 2EJF) superposed on EcBPL dimer (PDBID 2EWN)

The most recent BPL structure to be published is that of PhBPL (grey) in complex with BCCP (magenta). The arrow indicates the critical Lys residue which is biotinylated in the MKM motif. When this structure is superposed on EcBPL dimer (blue), the catalytic interface appears to be the same interface as that responsible for biotin synthesis repression in *E. coli.*

plasma membrane (46). There are five biotin-requiring carboxylases in yeast. The ACC1 gene encodes acetyl-CoA carboxylase that catalyses the formation of malonyl coA from acetyl coA in the cytoplasm (47) while the HFA1 gene encodes a mitochondrial isozyme (48). Likewise, PYC1 and PYC2 encode two isoforms of pyruvate carboxylase (49). The fifth enzyme is urea amidolyase which participates in two activities, one as a urea carboxylase and the other as an allophanate hydrolase. *S. cerevisiae* BPL (ScBPL), the product of a single gene locus, is required to biotinylate all five biotin-enzymes (47,49,50). Recently, a 45 kDa protein which has no detectable sequence homology to any biotin-dependent enzyme, has also been shown to incorporate biotin. Arc1p, a yeast amino-acyl t-RNA synthetase cofactor, that has no carboxylation activity however is indeed biotinylated by ScBPL (51). Interestingly, Arc1p does not contain the typical consensus motifs known for biotin domains (52). Notably, the extent of biotinylation of this protein increases with increased biotin in the yeast growth medium however biotinylation is not essential to the function of this enzyme (53). The mechanism by which ScBPL recognises and biotinylates this protein is thus far undetermined (51).

Structure of ScBPL

Although the 3-dimensional structure of this 76.4 kDa enzyme has yet to be determined, some structural cues in the enzyme have been reported. Using limited proteolysis experiments, it has been proposed that ScBPL contains a linker region (Lys240-Asn260) that acts as a hinge between the N-terminal 28 kDa region and the 50 kDa C-terminal region containing residues required for ligand binding. The linker region was made more resistant to proteolysis when complexed with biotin and ATP (54) suggesting conformational changes are

associated with ligand binding. Whether this region is truly a linker region with no role in the enzymatic function of the protein could be validated by the resolution of crystal structures of ScBPL in the presence and absence of biotin and ATP as well as by other protein analysis techniques such as mutagenesis.

Whilst the C-terminal region is necessary for biotin binding and catalysis, the role of the N-terminal region is poorly understood. However, the N-terminal may have novel functions as it has been identified as essential for catalysis. It has previously been shown that biotin-accepting domains can be recognised and biotinylated by BPLs from species other than its own (25,26,55). This observation facilitated isolation of the ScBPL gene by complementation of conditional lethal *birA1* strains of *E. coli* (56). Unlike full-length ScBPL, truncations of the N-terminal region (1-233, 1-369 and 1-409) were inactive in complementation assays (54). Furthermore, the truncations resulted in a greater than 3500- fold decrease in enzymatic activity in *in vitro* biotinylation assays relative to full-length ScBPL, highlighting the importance of the N-terminal domain for catalysis. An explanation for the molecular role of the N-terminal domain in catalysis is yet to be defined.

Importance of the N-terminal

The primary structure of the N-terminal region of yeast and human BPLs show no conservation to the N-terminal of BirA. It appears that these eukaryotic BPLs have become increasingly longer and complex through evolution (figure 4). Until this study, BLAST searches with the N-terminal sequence from both *S. cerevisiae* and human BPL had shown no homology to any protein sequences or plausible structures in available databanks (refer to chapter 6). Studies on both yeast and human BPLs have implicated a role for

Figure 4: Sequence alignment between bacteria, yeast and human BPL

Figure 4

Sequence alignment between species, as discussed in the text, shows that there is limited sequence homology. Residues that are completely conserved are indicated by (*), similar by (:) and limited similarity (.). Residues that are involved in binding biotinyl-5`-AMP in PhBPL and show similar conservation with EcBPL biotinol-5'-AMP are indicated by purple triangles. Why eukaryotes possess and require this large N-terminal extension, yet appear to have all the catalytic components and residues necessary for biotinylation still remains a mystery.

the N-terminal region in the catalytic mechanism (54,57). However the precise details by which this occurs is not completely understood. Furthermore, novel functions might also be associated with this region of the BPLs. It is envisaged that studying yeast BPL will aid in our understanding of other eukaryotic BPL, in particular human BPL. There are many yeast proteins with functions that are conserved in humans and for which comparisons have allowed the identification of a basic biological mechanism. One powerful tool that can be used to aid this overall understanding of proteins is structural determination by X-ray crystallography.

X-ray crystallography

One of the main objectives of this project is to solve the structures of BPLs from some of the species mentioned in this Introduction. X-ray crystallography has been a very valuable technique for the structural analysis of large proteins and their complexes. Unlike nuclear magnetic resonance spectroscopy, there is no size limit on the structures that can be studied. The first limiting step for crystallography is obtaining a homogeneous sample of the protein in question. This can usually be overcome by over-expression of recombinant proteins and the engineering of affinity tags to facilitate their purification. The next limiting step is optimisation of conditions required for crystal growth. There is no direct correlation between biological function, structure, size or physical properties of the protein and the correct conditions desired to produce diffraction quality crystals. Therefore, it is important to determine these conditions on a case-by-case basis for each protein (58).

Protein crystallisation methods

Nucleation and growth of crystals occurs in supersaturated solutions where the concentration of the protein exceeds its equilibrium solubility value. This is a function of the concentration of the macromolecule and parameters that affect its solubility. The higher saturation conditions (labile phase) support both growth and nucleation. Only the lower supersaturation conditions (metastable phase) support crystal growth (59).

Crystallisation via vapour diffusion

In this method, the unsaturated precipitant-containing protein is suspended over or in a well (known as the hanging or sitting drop method respectively) using a plate containing a reservoir of nucleation stimulating conditions. These conditions are usually set up in duplicate at room temperature and at 4^{o}C. Through vapour equilibration, crystal growth will occur during water exchange or evaporation from the protein drop to a more hygroscopic reservoir solution (59). Nucleation, the first sign of protein aggregation or crystallisation occurs by an increase in protein concentration through dehydration, thus driving the protein to “drop out” of solution as crystals. Solubility effects and physicochemical characteristics can be calculated for each protein so that the attraction between the molecules can be increased. For example, salts or neutral polymers can be used to stimulate attractive forces. Furthermore, the molecular weight, charge and pI of the protein can be used to calculate interaction potentials in solution as a function of environmental parameters, such as pH and suitable buffer.

Crystal growth

The formation of a crystal requires the ordered addition of 10^{15} molecules, occurring in a strictly identical and ordered fashion (59). This number leaves the opportunity for the misalignment of one or several molecules, molecular aggregation and effects of short and long term variability of the growth process. For example these include mutated, conformationally different protein species, vacancies in the lattice and trapped impurities that do not allow for useful diffraction patterns. Even though these point defects at the molecular level in the crystallisation process seem small, they can be replicated to subsequent layers during crystal growth, causing strain and leading to a “dead end” complex rather than a true structure (59).

Data collection

Once crystals have been identified, they are rapidly cooled and maintained at 100 ^{o}K during data collection. “Fast cooling” minimises radiation damage to the crystal and reduces background scatter and adsorption (60). This effectively increases the resolution limits of the technique and allows long-term storage and reuse of crystals.

X-ray storage-phosphor imaging-plate detector (IP):

The detector requirements are of fundamental importance for high quality diffraction pattern collection. The detector must have the following characteristics: a high detective quantum efficiency, wide dynamic range, linear and uniform response, high spatial resolution, large active area and high count-rate if used with synchrotron radiation (59). The X-ray storage-phosphor imaging-plate detector is suitable as it has the above performance

characteristics. The IP- detector is a flexible plastic plate with uniform coating of clusters of very small crystals of photostimulable phosphor and organic binder. This stores information as a function of the X-rays energy and when stimulated by visible light, emits a luminescence that is proportional to the absorbed X-ray intensity. Various computer readout systems can then calculate this pattern collected via the sum of various algorithms and matrices. The analysis from this data sums the contribution properties such as those from a single point atom of unit scattering, isotropy, probable density of atoms, expected variance and many more constraints and variables (59). CCD detectors are used for fast data collection and more efficient screening where photons pass through indium-tin-oxide to generate electron-hole pairs in the CCD. The CCD is then readout one row at a time to the serial shift register and the output amplified off the chip to an analogue-to-digital converter.

X-ray diffraction patterns

The first result from a diffractable crystal pattern after scaling and merging the data is an electron density map. The atomic model then serves as a representation of the electron density map generated from the knowledge of bond lengths and angles between amino acids. These are applied in the form of constraints and restraints during the refinement of the atomic model. When the resolution in the diffraction data increases, this improves the electron density and can correlate to individual atoms. Flexible regions in a structure may be resolved to < 1 Å yet be unresolved in an electron density map (59).

Limitations

Phases are lost during data collection, but there are two ways by which the phases can be deduced. Firstly, through isomorphous replacement where "heavy atoms" can be soaked into the crystal and perturb the diffraction pattern, which can then be used for positional information, where values for phase angles are found. Multi-wavelength anomalous diffraction (MAD) phasing can be used whereby a protein is produced as native and with the incorporation of selenomethionine. Another popular method is molecular replacement if there are similar structures available to the one being studied. Though there are many methods available to solve structures, the number of individual protein structures determined by X-ray crystallography is very small compared to the known number of protein sequences. This is due to the difficulties in producing the quantities of protein required for trials at complete homogeneity as well as the unpredictability of conditions for crystal growth. However the useful information that can come from producing protein structures from crystals can justify the time and effort taken to attain them.

Figures in the introduction were generated using PYMOL (61) and the PDB indicated in the figure title. Superposition was done using ssm function as part of coot (62). Figure 2c was generated using LIGPLOT (63). Alignment file (figure 4) was generated using ClustalW (64).

Aims of this project

Through this project I aim to expand upon the current database of known BPL structures. I will focus on the BPL from *Staphylococcus aureus* due to its clinical significance. This structure is imperative for structure-based drug design efforts, which will be firstly performed by *in silico* methods. I also intend to investigate eukaryotic BPLs, namely the enzymes from the model organism *Saccharomyces cerevisiae,* as well as the clinically important species *Candida albicans.* This will be the first study in which BPL from *C. albicans* has been characterised. I also wish to compare information gained from prokaryote BPL's and the yeast to that of HCS. Deciphering differences between BPLs and discovering the molecular mechanism will avoid the targeting of novel therapeutics to the human enzyme in our pursuit of novel antibiotics and antifungals.

Chapter 2

Distant relatives of Biotin Protein Ligase aid in understanding multiple carboxylase deficiency.

This chapter is an extensive literature review on what is known about the human form of Biotin Protein Ligase, also known as Holocarboxylase Synthetase (HCS) as well as mutations in this enzyme that lead to the disease Multiple Carboxylase Deficiency (MCD). Various expression systems, buffers, lysis techniques and protease inhibitors were trialed to attain full length HCS in order to solve the 3-dimensional structure by X-ray crystallography. Although my attempts were unsuccessful, sequence homology of the catalytic domain of HCS combined with structures available in the PDB allowed a 3-dimensional structure to be modeled and position of mutations causing MCD highlighted providing a rationale as to why these mutated residues lead to disease states.

Chapter 2

Distant relatives of Biotin Protein Ligase aid in understanding multiple carboxylase deficiency.

Author	Contribution	Signature
[+]Nicole R. Pendini	Contributed approximately 1/3 of the written text, primary on the structure and function of biotin-enzymes and BPLs. Developed model of HCS, and used this to predict the effect of mutation on enzyme structure and function (Table 2). Provided Figures 3 and 4. Collated references and entire manuscript.	
[+]Lisa M. Bailey	Contributed approximately 1/3 of the written text, primary on sections pertaining to MCD. Collated published data for Table 1.	
Grant W. Booker	Assisted in proof reading and editing of the manuscript.	
Matthew C. Wilce	Assisted in development of the HCS model and analysis of mutation data from MCD patients.	
John C. Wallace	Assisted in proofreading and editing of the manuscript.	
Steven W. Polyak	Provided approximately 1/3 of the written text. Provided figures 1 and 2. Primary role in collation of the manuscript and editing of revisions.	

[+] = Equal first authorship

ELSEVIER LICENSE
TERMS AND CONDITIONS

Aug 19, 2008

This is a License Agreement between Nicole R Pendini ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Nicole R Pendini
Customer address	biochemistry, bld 13D, wellington rd Victoria, other 3800
License Number	2012780660657
License date	Aug 19, 2008
Licensed content publisher	Elsevier
Licensed content publication	Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics
Licensed content title	Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency
Licensed content author	Nicole R. Pendini, Lisa M. Bailey, Grant W. Booker, Matthew C. Wilce, John C. Wallace and Steven W. Polyak
Licensed content date	July-August 2008
Volume number	1784
Issue number	7-8
Pages	10
Type of Use	Thesis / Dissertation
Portion	Full article
Format	Electronic
You are an author of the Elsevier article	Yes
Are you translating?	No
Purchase order number	
Expected publication date	Jul 2008
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
Value added tax 0.0%	0.00 USD

Biochimica et Biophysica Acta 1784 (2008) 973–982

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Review

Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency

Nicole R. Pendini [a,1], Lisa M. Bailey [a,1,2], Grant W. Booker [a], Matthew C. Wilce [b], John C. Wallace [a], Steven W. Polyak [a,*]

[a] School of Molecular and Biomedical Science, University of Adelaide, North Tce, Adelaide, South Australia 5005, Australia
[b] Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria, Australia

ARTICLE INFO

Article history:
Received 16 January 2008
Received in revised form 16 March 2008
Accepted 26 March 2008
Available online 9 April 2008

Keywords:
Biotin protein ligase
Holocarboxylase synthetase
HCS
Multiple carboxylase deficiency
Protein structure and function

ABSTRACT

The attachment of biotin onto the biotin-dependent enzymes is catalysed by biotin protein ligase (BPL), also known as holocarboxylase synthase HCS in mammals. Mammals contain five biotin-enzymes that participate in a number of important metabolic pathways such as fatty acid biogenesis, gluconeogenesis and amino acid catabolism. All mammalian biotin-enzymes are post-translationally biotinylated, and therefore activated, through the action of a single HCS. Substrate recognition by BPLs occurs through conserved structural cues that govern the specificity of biotinylation. Defects in biotin metabolism, including HCS, give rise to multiple carboxylase deficiency (MCD). Here we review the literature on this important enzyme. In particular, we focus on the new information that has been learned about BPL's from a number of recently published protein structures. Through molecular modelling studies insights into the structural basis of HCS deficiency in MCD are discussed.

1. Biotin is an enzyme cofactor

Biotin, also known as vitamin H or B_7, is an essential micronutrient in mammals. Biotin was first discovered by Kögl and Tönnis in 1934 while studying the growth requirements of yeast in synthetic media [1]. Mammals are auxotrophic for biotin and obtain it through dietary sources and from intestinal bacteria that can produce biotin *de novo* [2]. Biotin is actively absorbed into the circulation, and enters cells via transporter proteins. Uptake studies in placenta [3], brain [4] and intestinal tract [5] have been reviewed recently. In most mammalian tissues the primary transport protein is the sodium-dependent multivitamin transporter, SMVT1, which has been cloned from rat [6] and human [7] placenta. Alternate biotin transporters, such as the monocarboxylate transporter MCT1 and solute carrier family 19 member 3 SLC19A3, have been also been proposed in white cells and keratinocytes [8,9]. Within the cell, the most extensively characterised and understood function for biotin is to serve as a cofactor for biotin-dependent carboxylases where it is required for the transfer of carbon dioxide from bicarbonate to organic acid metabolites [10–13]. Biotin is covalently attached to this family of enzymes by BPL. As the biotin-dependent enzymes are turned over, biotin is hydrolysed off the degraded peptide fragments by biotinidase, allowing recycling of the vitamin.

The five mammalian biotin-dependent enzymes are all functional carboxylases and are essential due to their involvement in fatty acid synthesis, gluconeogenesis and amino acid catabolism [14]. Acetyl CoA Carboxylase-1 (265 kDa) is located in the cytoplasm where it performs the essential role of converting acetyl CoA to malonyl CoA, the first committed step in fatty acid synthesis, required for membrane biogenesis. The enzyme exists either as catalytic homodimers or associated with more highly active filamentous fibres [15]. Acetyl CoA Carboxylase-2 (280 kDa) is localised to the outer membrane of the mitochondria where, like acetyl CoA carboxylase-1, it catalyses the formation of malonyl CoA [16]. This pool of malonyl CoA regulates the movement of fatty acids into the mitochondria for β-oxidation by functioning as an allosteric inhibitor of carnitine palmitoyl CoA transferase [17]. The three remaining mammalian biotin-dependent enzymes all reside within the mitochondria. Pyruvate Carboxylase forms a native tetrameric enzyme of 130 kDa identical subunits. It catalyses the formation of oxaloacetate from pyruvate. Oxaloacetate is required for gluconeogenesis in the liver and to replenish tricarboxylic acid cycle intermediates used for biosynthetic purposes such as citrate for fatty acid synthesis and for glutamate synthesis as a neurotransmitter [18]. The first structure for this enzyme in its tetrameric form was recently resolved by X-ray crystallography [19]. Propionyl CoA Carboxylase catalyses the generation of methylmalonyl CoA from

* Corresponding author. Molecular Lifesciences Building, University of Adelaide, North Tce, Adelaide, South Australia 5005, Australia. Tel.: +61 8 8303 5289; fax: +61 8 8303 4362.
E-mail address: steven.polyak@adelaide.edu.au (S.W. Polyak).
[1] Equal authorship.
[2] Current address: The Royal Institution of Great Britain, 21 Albemarle Street, London W1S 4BS, UK.

1570-9639/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2008.03.011

propionyl CoA, derived from metabolism of branched amino acids such as valine, isoleucine, methionine and threonine as well as odd-numbered fatty acid chains [20]. It consists of a $\alpha 6\beta 6$ heteropolymer with biotin attached to the α-subunits. Likewise, 3-Methylcrotonyl CoA carboxylase is also a $\alpha 6\beta 6$ dodecameric enzyme composed of biotin-containing α-subunits together with β-subunits. This enzyme catalyses carboxylation of 3-methylcrotonyl CoA to 3-methylglutaconyl CoA during the catabolism of leucine [21]. A deficiency in this pathway results in elevated excretion of 3-hydroxyisovaleric acid in the urine, an observation that has been exploited as a reliable indicator of biotin-deficiency in patients [22].

All biotin-dependent enzymes utilise the enzyme-bound biotin group for the transfer of CO_2 between metabolites. As such all enzymes in this family have three distinct domains that are required for catalysis; the biotin domain, the biotin carboxylase domain (first partial reaction site) and the carboxyl transferase domain (second partial reaction site) [12,23]. During its lifecycle the biotin domain makes three heterologous protein:protein interactions. First, biotin is covalently attached onto one specific lysine within the biotin domain by BPL. The biotinylated domain then oscillates between the two partial reaction sites within the biotin-dependent enzyme. This swinging arm mechanism is conserved amongst all classes of biotin-dependent enzymes. In the case of the mammalian biotin-dependent carboxylases, enzyme-bound biotin is first carboxylated by CO_2 in the form of dissolved bicarbonate in an ATP-dependent reaction (reaction 1) [11,13,24]. This first partial reaction requires cooperation of the biotin carboxylase and biotin domains whereby the N_1 of biotin is carboxylated through a carboxy-phosphate intermediate requiring ATP and bicarbonate. The second partial reaction then proceeds by the transfer of the carboxyl anion from carboxybiotin onto the appropriate acceptor substrate (reaction 2). It has been proposed that the decarboxylation of carboxybiotin proceeds along a pathway where CO_2 and the enolate of biotin are formed [13]. Following protonation, carboxylated biotin can again be formed at the first partial reaction site under physiological conditions.

$$\text{Enzyme–biotin} + \text{MgATP} + HCO_3^- \leftrightarrow \text{MgADP} + P_i + \text{Enz–biotin–}CO_2 \quad (1)$$

$$\underset{\text{Acceptor metabolite}}{\text{Enz–biotin–}CO_2 + \text{R–H}} \leftrightarrow \underset{\text{Carboxylated metabolite}}{\text{Enz–biotin}} + \text{R–}CO_2 \quad (2)$$

2. Biotinylation

HCS is the enzyme responsible for specifically attaching biotin onto the mammalian biotin domains. Biotinylation is catalysed through a two-step reaction where biotin is first activated to biotinyl-5′-AMP in an ATP-dependent manner [25]. The biotin is then transferred onto the ε-amino group of a specific target lysine residue (outlined in reactions 3 and 4). The reaction mechanism is related to that of amino acyl-tRNA synthetases [26] and lipoyl ligases [27] where the reaction proceeds through the formation of an adenylated intermediate, suggesting a common ancestral relationship. Furthermore the protein:protein interaction between BPL and biotin-dependent enzymes is also highly conserved as it has been shown that BPL's are interchangeable between organisms [28–32] although BPLs appear to have the highest affinity for their natural substrate [33]. Cross-species activity of BPLs permitted functional cloning of the cDNA for human [34] and yeast BPLs [35] through genetic complementation with a conditional lethal *birA*$^-$ strain of *E. coli*. This interchangeable interaction is possible due to structural conservation of both the BPLs and biotin domains. Homology at the primary structure level is observed not only between all biotin-dependent enzymes but also with non-biotin requiring enzymes such as the lipoyl enzymes [27]. Despite this, aberrant biotinylation or lipoylation of inappropriate substrates is avoided due to structural cues in both enzyme and substrate that will be discussed in further detail.

$$\text{biotin} + \text{ATP} \leftrightarrow \text{biotinyl} - 5' - \text{AMP} + PP_i \quad (3)$$

$$\text{biotinyl} - 5' - \text{AMP} + \text{apo} - \text{protein} \rightarrow \text{biotinylated holo} - \text{protein} + \text{AMP} \quad (4)$$

3. Structure of the biotin domain

The structures of the biotin domains from several species have been reported and deposited in the protein database. The first and one of the most fully characterised structures is from the *Escherichia coli* biotin carboxyl carrier protein (BCCP), a subunit of the acetyl CoA carboxylase complex. The C-terminal 87 amino acids of BCCP (BCCP-87) is recognised and biotinylated by BPL as well as full-length BCCP [36]. This is due to the fact that this peptide adopts the tertiary structure that is required for the interaction with BPL. The BCCP-87 structure has been determined by both X-ray crystallography [37] and NMR spectroscopy [38] with both techniques giving essentially identical structures. The apo (unbiotinylated) and holo (biotinylated) forms of the BCCP-87 are similar with only subtle and localised structural changes observed during NMR analysis [38]. Furthermore, the structures of other biotin domains can also be superimposed on BCCP-87 (see Fig. 1) highlighting a conserved, signature structural motif present in all biotin-dependent enzymes. This protein fold consists of two anti-parallel β-strands that form a β-sandwich fold. The domain is stabilised by a core of conserved glycine and hydrophobic amino acids that impair biotinylation if mutated or deleted by truncation [28,33,39]. Thus, the recognition of a biotin domain by BPL firstly relies on the substrate adopting an appropriate tertiary structure.

The biotinylated lysine itself is located in a hairpin turn between β-strands 4 and 5 in the centre of the polypeptide, present in a Met–Lys–Met motif that is essentially invariant in all biotin domains [40]. The flanking methionine residues have been proposed to aid in biotin-carboxylation during the life cycle of the biotin-dependent enzymes [41,42]. Unlike other forms of post-translational modification, the sequence motif surrounding the target lysine is not the principal determinant for modification. Indeed, it is the precise positioning of the lysine residue within the context of the structured biotin domain that is important. Movement of the lysine by even one position around the loop abolishes biotinylation [43]. Whilst the flanking residues in the loop are not essential for biotinylation they have been shown to affect affinity for BPL. Mutational analysis by both directed [29,39,43] and random approaches [33] has revealed that only conservative substitutions permit biotinylation. Interestingly, the lipoyl-accepting proteins adopt a conformation analogous to the biotin domains. Here the lysine residue targeted for lipoylation is also present in a prominent hairpin loop but occurs in a conserved Asp–Lys–Ala motif. Replacement of this motif with the biotinylation motif Met–Lys–Met in the lipoyl domain of *Bacillus stearothermophilus* pyruvate dehydrogenase complex did not permit biotinylation [44] again highlighting that the Met–Lys–Met motif alone is not sufficient for appropriate modification by the BPL.

Some bacterial domains, including BCCP-87, contain a "thumb"-like protrusion between the β2 and β3 strands that is important for several reasons. Firstly, the thumb makes contact with the biotinyl group partially burying the cofactor in the surface of the protein [37,38]. It has been proposed that this thumb may function as a mobile lid for either, or possibly both, the biotin carboxylase or carboxyl-transferase active sites in the biotin-dependent enzyme [45]. The function of this lid could aid to prevent solvation of the active sites, thereby aiding in the transfer of CO_2 from carboxybiotin to acetyl CoA. Secondly, the thumb is required for dimerisation of BCCP necessary for the formation of the active acetyl CoA carboxylase complex [45].

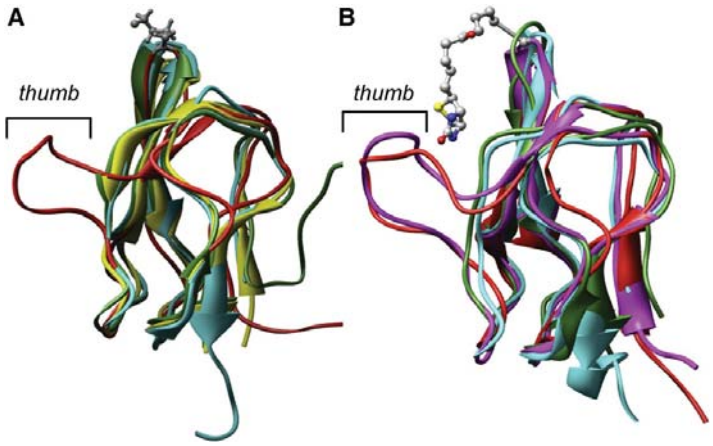


Fig. 1. Structural conservation of the biotin-accepting domain. Co-ordinates deposited in the PDB were downloaded and superimposed on other biotin domains to demonstrate the structural conservation across species. The structures of (A) apo and (B) holo biotin domains are shown in ribbon representation. The target lysine side chain, either free or covalently bonded to biotin, is represented in ball and stick mode located in an exposed hairpin loop found between β-strands 4 and 5. The position of the "thumb" insertion in *E. coli* BCCP is indicated on the structure. A. The apo-structures of *E. coli* BCCP (red— 3BDO) [38], *Propionibacterium shermanii* transcarboxylase (green— 1DCZ) [109], *Bacillus subtilis* BCCP (cyan— 1Z7T) [48], and *Pyrococcus horikoshii* BCCP (yellow— 2D5D) [79] are included. B. Biotinylated domains from *E. coli* BCCP, determined by both X-ray crystallography (magenta— 1BDO) [37] and NMR (red— 2BDO) [38], *P. shermanii* transcarboxylase (green— 1O78) [47] and *B. subtilis* BCCP (1Z6H) [48] are overlaid. Numbers in brackets are the PDB assignment to each structure.

Finally, the thumb functions to inhibit the aberrant lipoylation of the target lysine by lipoyl protein ligase [46]. Removal of the amino acid residues that constitute the thumb by mutagenesis rendered BCCP-87 a favourable substrate for lipoylation but abolished biotinylation [46].

The thumb structure, however, is not a highly conserved feature amongst all biotin domains (Fig. 1). Many, including those from all five mammalian biotin-dependent enzymes, do not contain this insertion. Interestingly, it appears the contact between biotin and protein might be a conserved feature and important for catalysis as similar contacts have been observed in the "thumbless" domains from *P. shermanii* transcarboxylase [47] and the biotinyl/lipoyl attachment protein of *B. subtilis* [48]. The significance of this requires further investigation. Furthermore, how aberrant lipoylation and biotinylation is avoided in mammalian cells containing multiple targets is not easily explained for the thumbless substrates. It is likely that the modifying enzymes themselves have evolved proof reading functions to ensure precise substrate selection.

4. Human holocarboxylase synthetase

A high-resolution structure of human HCS has not yet been reported. However, from truncation and mutagenesis studies as well as amino acid alignments, we can deduce some important structural features of this enzyme. The protein is predicted to contain three structural domains, one in the N-terminal half (amino acids 166–290) and two in the C-terminal half (amino acids 460–669 and 670–726) (Fig. 2). Deletion analysis by Campeau and Gravel (2001) demonstrated that the minimal functional protein consists of the C-terminal 349 residues. This region of the enzyme contains the highest degree of conservation across all species [49,50] and is likely to adopt the conserved protein structure required for the cross-species biotinylation of biotin domains that has been reported [28,30–32]. Furthermore, this structural topology positions BPL in the same family as the lipoyl ligases [27] and amino acyl-tRNA synthetases [26]. In this review we discuss the features of the catalytic domain in further detail and its relevance to human disease.

Whilst the function of the C-terminal half of the protein is understood, the N-terminal half has not been as well explored. Although the N-terminal region is not essential for activity it has been implicated in catalysis [49]. HCS with an N-terminal deletion of the first 165 residues was able to rescue BPL-deficient *E. coli*, but was unstable and not detected by Western blotting. Of particular significance were deletions up to Ala235 or Thr266 that compromised activity, whereas further deletions up to Lys378 restored activity. This was also observed by Sakamoto et al. (1999) who observed that deletion of residues 1–234 almost abolished HCS activity in SV40-transformed fibroblasts [51]. These studies indicate that the region between Leu166 and Arg290 is required for catalysis, possibly through an interaction with the catalytic domain. A mechanism where the N-terminal region interacts with the catalytic domain was first proposed for yeast BPL [52] that also contains a large N-terminal extension required for activity. Domain mapping of HCS by limited proteolysis supports the proposal of a structured domain encompassing Leu166 to Arg290 (Swift, Bailey, Polyak and Wallace, unpublished data). Additionally, the Campeau and Gravel study provided *in vivo* data showing HCS could indeed discriminate between the endogenous bacterial BCCP and the biotin domain of human propionyl CoA carboxylase [49]. These data imply HCS may well contain proofreading activity to select appropriate substrates, but further work is required to understand this in greater detail.

5. Multiple carboxylase deficiency

Multiple carboxylase deficiency (MCD) is the condition arising from the lack of activity of the biotin-dependent enzymes. It can be caused by a defect in biotinylation of carboxylases (HCS deficiency, OMIM 253270), a deficit in recycling of biotin in the cell (biotinidase deficiency, OMIM 253260) or failure of biotin transport. Whilst there are numerous clinically and biochemically common features, individual cases are extremely variable. Symptoms include ketoacidosis, feeding difficulties, hypotonia, seizures, developmental delay and dermal abnormalities such as rashes, dryness of the skin and alopecia.

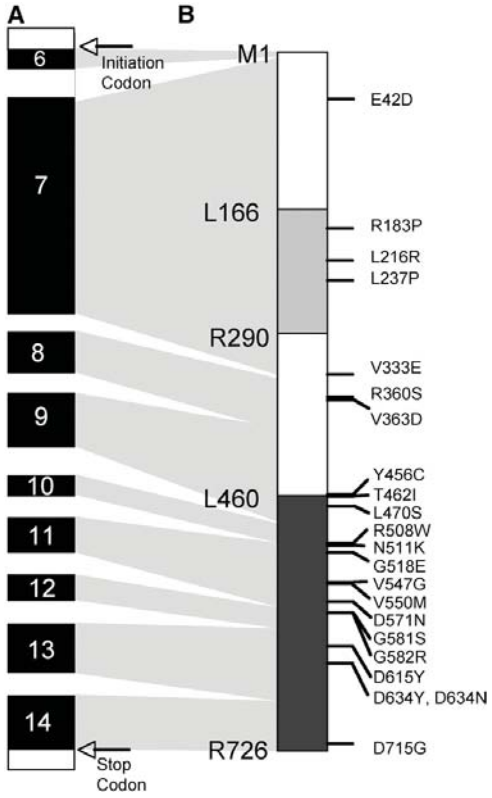


Fig. 2. Schematic representation of the HCS gene and protein. A. The gene structure of HCS is shown with the black boxes representing exons encoding the protein. Each exon is numbered and the positions of the initiation and termination codons shown. The relationship between the exons and the protein is represented by grey shading between the two schematic diagrams. B. Representation of the HCS protein, highlighting the conserved C-terminal catalytic region (dark grey box) and the domain in the N-terminal extension required for catalysis (light grey box). Positions of mutations arising in MCD are indicated on the right. Figure adapted from [59].

In severe or untreated cases this can lead to coma and death [53]. This review will focus upon HCS deficiency in MCD.

Saunders and co-workers first implicated a defect in HCS in patients with MCD [54]. Activities of the three mitochondrial biotin-dependent enzymes, propionyl coA carboxylase, methylcrotonyl coA carboxylase and pyruvate carboxylase, were all assayed in fibroblasts from two patients presenting with biotin-responsive organicacidemia. It was observed that activities of all 3 enzymes were severely reduced in low biotin media but could be rescued by addition of biotin-rich media, suggesting a common requirement for HCS. Involvement of HCS was confirmed by Burri et al. [55] who developed an HCS assay using apo-PCC partially purified from the livers of biotin-deficient rats. In this assay, the K_M for biotin was elevated 60-fold in extracts from patient fibroblasts relative to normal controls. This study was the first to demonstrate HCS deficiency in an MCD patient.

Subsequently it was shown that defects in biotinidase could also give rise to MCD [56]. Initially, age of onset was used to define the defect causing MCD. HCS deficiency usually presents with ketoacidosis in neonates within the first few days of life, and is associated with a poor prognosis if left untreated. Biotinidase deficiency generally presents with a gradual onset of symptoms including audio-visual pathologies (which are not associated with HCS deficiency), limb muscle weakness and progressive neurologic abnormalities [53]. Additionally, biotinidase deficiency does not present until initial biotin stores are depleted, and is also dependent on the amount of biotin in the diet and residual biotinidase activity. Generally, biotinidase deficiency presents at a later age (average onset at 3 months) [57]. However, HCS deficiency has been known to present as late as 8 years [58], hence the clinical relevance of age of onset has decreased as a diagnostic tool with diagnosis relying instead on either DNA assay for known mutations [59] or enzyme activity assay. Neonatal screening by tandem mass spectroscopy for HCS deficiency is now available using filter paper dried blood spot samples [60], and pre-natal molecular diagnosis by DNA analysis can be achieved through chorionic villus sampling at 11 weeks gestation in families with a history of the disease [61].

There have been multiple reported mutations in the HCS gene that give rise to MCD (recently reviewed in Suzuki et al., 2005 [59]; summarised in Fig. 2 and Table 1). The severity of disease varies greatly with the genotype of each individual. Broadly, the mutations can be classified into two groups. Firstly the K_M mutants are amino acid substitutions that result in an enzyme with decreased affinity for biotin. Patients bearing these mutant HCS respond well to oral administration of pharmacological doses of biotin. Secondly, the V_{MAX} mutants are substitutions where the activity of the enzyme is compromised but not restored with higher concentrations of biotin. Responsiveness to biotin therapy varies for patients with these mutations. Below is an overview of these HCS mutations in the context of the BPL structure. To date there have been no reports of an X-ray structure for human HCS. However recent studies reporting new structures of BPL from bacteria *E. coli* and archea *P. horikoshii* OT3 have provided powerful insights into the structural basis for HCS deficiency in MCD.

6. *E. coli* BPL

The most characterised of the BPL's is that of *E. coli* (*Ec*BPL), also known as the biotin-inducible repressor BirA. *Ec*BPL is a bifunctional molecule of 35.5 kDa with the ability to carry out two roles: 1) to catalyse biotin attachment onto BCCP and 2) function as a transcriptional repressor of the biotin biosynthetic operon (reviewed [62]). *Ec*BPL has the capability of recognising and biotinylating only one protein, namely BCCP, out of over 4000 proteins in the *E. coli* proteome. Furthermore, biotin is only ever attached to the side chain of the single target lysine required for activity of acetyl CoA carboxylase [29,39].

*Ec*BPL contains three distinct domains that have been determined through X-ray crystallography, at 2.3 Å resolution. The monomeric structure measures 75 Å×35 Å×30 Å. The N-terminal 22–46 residues adopt a helix–turn–helix motif, a structure associated with DNA binding proteins [63]. The central domain is conserved amongst all BPLs and contains the binding sites for ligands biotin and ATP. The central domain consists of five α-helices, 7 strands of mixed β-sheets as well as four poorly defined loops that appear in pairs in the 3D structure. These loops, consisting of residues 110–128, 212–233, 140–146 and 193–199, are associated with conformational changes to *Ec*BPL upon ligand binding. The C-terminal domain consists of 6 strands that form a β-sandwich that seals the end of the enzyme and has been found to function in the transfer of biotin onto BCCP [63–65].

The crystal structure of *Ec*BPL in complex with biotin has also been resolved [66]. Monomeric *Ec*BPL crystals were destroyed when soaked with biocytin, suggesting that conformational changes occurred upon biotin binding [63]. It was previously thought that residues ^{115}GRGRRG^{120}, found in the partially unstructured surface loop composed of residues 110–128 [63], bound ATP as it is akin to the GXGXXG motif found in many nucleotide-binding sequences in proteins [63,67,68]. Mutational analysis of this sequence resulted in

Table 1
Biochemical characterisation of HCS missense mutations reported in MCD patients

Mutation	Activity (% wildtype)	K_M Biotin (fold over wildtype)	References
Wildtype	100	1	[89]
E42D	120	N.D.	[113]
R183P	1.7	0.6	[51]
L216R	0.3	1.4	[51]
L237P	1.2–4.3	0.4–1.2	[51,88,89]
V333E	2–10	1.5	[51,88,89]
R360S	22.0	N.D.	[113]
V363D	3.7	1.1	[51]
Y456C	0.2	N.D.	[113]
T462I	<10	N.D.	[88]
L470S	4.3	N.D.	[113]
R508W	34.5*	23*	[82]
V547G	3.4	N.D.	[113]
V550M	16.6	6.5	[90]
D571N	0.1	N.D.	[88]
G581S	<10	44.3	[51,88,89]
D634Y	12.0	N.D.	[113]

Amino acid substitutions in the catalytic region (residues 422–726) induce a 3–44-fold elevated K_M for biotin compared to wildtype HCS. N.D Not determined.

defects in enzymatic function [69], highlighting the importance of this region. However, kinetic experiments with these mutant enzymes as well as wild type *Ec*BPL showed a similar K_M for ATP, suggesting this region does not have a role in ATP binding. Nevertheless, the binding of biotin or biotinyl-5′-AMP to two of the *Ec*BPL mutants was compromised [69]. Furthermore, when crystals were grown in the presence of biotin, the loop consisting of residues 110–128 become ordered and encompassed the biotin cofactor [65]. More recently the structure of *Ec*BPL with the co-repressor analogue biotinol-5′-AMP has been resolved to 2.8 Å [70] and further highlights the importance of the loops. Residues Arg118 and Arg121 that reside within the ^{115}GRGRRGR^{121} motif make critical phosphate binding interactions where as Arg116 and Gly117 interact with biotin exclusively [70]. Residue Arg119 is largely solvent-exposed having little effect on ligand binding. The loop consisting of residues 212–233 folds over the purine ring of the adenylate [70]. Here, a cluster of hydrophobic amino acids assemble over the adenine base and inhibit dissociation of biotinyl-5′-AMP [71].

The helix–turn–helix N-terminal domain of *Ec*BPL is not directly involved in the catalytic mechanism. Experiments on truncated *Ec*BPL variants showed synthesis of biotinyl-5′-AMP and biotinyl-transfer were not compromised. However, these deletions showed a significant decrease in the affinity for biotin and biotinyl-5′-AMP [72]. A 100-fold increase in the rate of dissociation of the *Ec*BPL–biotinyl-5′-AMP complex was observed for the truncated proteins compared to full-length *Ec*BPL [72]. The N-terminal region of *Ec*BPL is essential for DNA binding and regulation of the biotin biosynthetic operon. In *E. coli*, the *bio* (ABFCD) locus contains five of the six genes required for biotin biosynthesis [73]. The sixth gene (*bioH*) and the *birA* gene, encoding *Ec*BPL, are located on separate regions of the *E. coli* chromosome. Synthesis of biotin is repressed by exogenous biotin concentrations greater than 40 nM and partially repressed by endogenous biotin synthesis [74]. Two *Ec*BPL monomers in complex with biotinyl-5′-AMP bind cooperatively with the biotin operator [75] to activate dimerisation and DNA binding [76]. The co-repressor increases the stability of the *Ec*BPL dimer by 3.7 kcal/mol compared to biotin alone [76]. In the crystal structure of the *Ec*BPL biotinol-AMP complex (PDB: 2EWN) the ligand binding surface loops that are involved in the dimer interface are more ordered, leading to a 12° bend at the dimer interface compared to *Ec*BPL plus biotin (PDB: 1HXD). As the two holo*Ec*BPL monomers bind 12-bp at the termini of the 40-bp *bioO* operator sequence, it is believed that the DNA bends slightly in order to accommodate specific interactions with one another. This interaction represses the initiation of transcription of the biotin biosynthetic operon by blocking the binding of RNA polymerases to the two promoters Pa and Pb [77]. When apo-BCCP accumulates in the cell it is able to compete with the *Ec*BPL dimerisation thereby inducing transcriptional derepression. Alternatively, in the absence of apo-BCCP, holo*Ec*BPL will accumulate and dimerise thus facilitating transcriptional repression. Therefore, alterations to the intracellular concentrations of apo-BCCP and biotin regulate the functional status of *Ec*BPL. Further information about *Ec*BPL and biotin operon regulation in *E. coli* is thoroughly summarised in the excellent review by Beckett [62].

7. BPL from *P. horikoshii* OT3

Until recently, the only structural information about BPL's came from that of *Ec*BPL. However, the structure of BPL from *P. horikoshii* OT3 (an archaeal hyper thermophile) has provided further insights into BPL function. *Ph*BPL in complex with biotin, ATP and biotinyl-5′-AMP have been resolved at 1.6 Á, 1.6 Á, 1.45 Á respectively [78]. Unlike *Ec*BPL, *Ph*BPL does not contain a DNA binding motif so does not function as a transcriptional repressor. Also the protein is a constitutive dimer (ie dimerises independently of biotin binding) that is structurally distinct from the holo*Ec*BPL dimer due to different dimerisation interfaces [78]. Notwithstanding, the domains required for catalysis are structurally similar with the two structures showing an r.m.s.d of 2.41 Á when C^{α} atoms are superimposed [78]. Furthermore, the primary structures of the catalytic domains are conserved (31% sequence identity), especially in amino acids required for ligand binding and catalysis. As observed with *Ec*BPL, biotin is anchored into the protein through hydrogen bonds with the hydrophilic interior of the biotin-binding pocket whilst the hydrophobic tail is stabilised by interactions with hydrophobic residues. Importantly, the *Ph*BPL structure provided the first structural insights into nucleotide binding. The ATP binding site lies adjacent to the biotin site and is more surface-exposed in the absence of nucleotide. Like *Ec*BPL, several disordered loops in the free enzyme become more ordered upon ligand binding. Once in complex, the adenosine phosphate and biotin carboxyl groups are orientated to facilitate adenylation (ie the first partial reaction). Overlays of the biotin- and ADP-complexed structures with the biotinyl-5′-AMP complex showed that the reaction intermediate binds in essentially the same position as its two constituent moieties, thereby minimising the protein conformational changes required for adenylation. An invariant lysine residue, (*Ph*BPL Lys111, *Ec*BPL Lys183) juxtaposed in the active site, has been proposed to play an important catalytic role during the synthesis of biotinyl-5′-AMP. Here the positive charge provided by the lysine side chain withdraws an electron from the nucleotide phosphate thereby facilitating nucleophilic attack by the carboxylic group of biotin. The reaction mechanism is currently being analysed through additional structural studies. These include co-crystallisation of the enzyme with a biotin domain necessary for the second partial reaction [79]. Though little is known about the biology of *Ph*BPL in *P. horikoshii* OT3, these high-resolution studies have provided invaluable insights into BPL.

8. Molecular analysis of mutations causing HCS deficiency

In order to understand the molecular mechanisms by which the HCS deficiency mutations affect HCS, a molecular model of the C-terminal catalytic region was prepared. HCS shares 23% sequence identity with *Ph*BPL, which is higher than that observed for *Ec*BPL (21%). Therefore, the three-dimensional co-ordinates of the catalytic domain of HCS were derived by homology modeling using the *Ph*BPL structure as a template (PDB ID: 1WQW) [78] (Fig. 3). This was performed using programs from Schrodinger, LLC [80]. Surface loops of HCS, which varied significantly in sequence to the *Ph*BPL structure, were either homology modeled using *Ec*BPL where appropriate or fragments from structures within the PDB were best fitted

using the modeling programme "O" [81]. This uses information from high-resolution structures in the PDB, geometrical and stereochemical dictionaries and validation tools such as Ramachandran analysis. The positions of mutations causing HCS deficiency have been mapped onto *Ph*BPL as well as the HCS model (Fig. 4). Discussion of the potential impact of such mutations on HCS structure and function is described below.

9. K_M mutants

There are many reports of MCD patients who have a variant HCS with an elevated K_M for biotin, although the estimated K_M values obtained vary between 4 and 260 nM depending on the assay system used [51,58,82–85]. In the cases where the mutation results in an elevated K_M for biotin, treatment with oral doses of biotin up to 10 mg/day is usually sufficient to resolve symptoms as the increase in biotin concentration is enough to overcome the decreased affinity. Treatment results in sufficient HCS activity to activate the biotin-carboxylases to normal levels.

Most of the mutations that result in increased K_M for biotin lie in the C-terminal catalytic portion of the enzyme. These mutations have been modelled onto the HCS structure (Fig. 4) and their functional importance summarised (Table 2). The model highlights those amino acid substitutions that cluster around the biotin and ATP sites. Notable is Arg508 (Arg118 in *Ec*BPL, Arg48 in *Ph*BPL), a highly conserved residue in all BPL's. In the *E. coli* homologue it is located on the flexible loop encompassing amino acids 110–128 that become structured upon ligand binding. Mutation of Arg118 to glycine in *Ec*BPL caused increased dissociation rates for biotin and biotinyl-5′-AMP [86], and consequently, promiscuous biotinylation of inappropriate proteins [87]. It is clear that a mutation to this residue results in an altered affinity of BPL for biotin. Interestingly, Asn511 (Arg121 in *Ec*BPL, Arg51 in *Ph*BPL) [88] also resides in the flexible loop with Arg508. Also of note is Asp571 that occurs at a very highly conserved position within the signature KWPND motif present in all BPLs. This mutation does not make direct contact with either ligand but is critical for the positioning of Lys579 (Lys183 in *Ec*BPL, Lys111 in *Ph*BPL) in the AMP binding pocket. Conserved glycine residues that lie in (Gly518) or near (Gly581 and Gly582) the biotin-binding pocket are sites for mutation, as is Asp615 in HCS (Asp220 in *Ec*BPL Asp138 in *Ph*BPL) in the nucleotide site. Other mutations are not as easily explained but most likely result in subtle conformational changes that influence ligand binding or catalysis.

10. V_{MAX} mutants

The V_{MAX} group of mutations, namely Arg183→Pro, Leu216→Arg and Leu237→Pro, map outside of the conserved catalytic core of HCS to the N-terminal region. Patients with these mutant HCS alleles vary in their responsiveness to biotin treatment. Although the N-terminal region has been implicated in catalysis [49] the precise role played by this extension is poorly understood. These mutations have provided an insight into the structure and function relationships of HCS. Interestingly, alignment of the BPL sequences from human, dog, horse, mouse, rat, chicken, *Danio* and *Fugu* show that the N-terminal extension is conserved amongst these vertebrate species. Furthermore, the equivalent amino acids at the above positions are invariant between these species, again highlighting their importance.

The Leu237→Pro mutation was the first mutation outside of the catalytic domain reported [89]. This enzyme, when recombinantly expressed, has severely reduced biotinylation activity [51,90]. For one patient, homozygous for the mutation, treatment with 40 mg biotin/day resolved all clinical and biochemical abnormalities, and development was normal at 9 months. However, follow up at 5 years showed a lowered IQ, for this patient who was otherwise symptom-free. Another patient required biotin treatment at 80 mg/day to control dermatological symptoms and showed delays in mental development at age 7 [89]. Yang et al. [91] reported two patients heterozygous for Leu237→Pro, with the other allele containing a truncated, non-functional protein.

Fig. 3. Model of the HCS catalytic domain. A model of HCS was generated with the X-ray co-ordinates for *Ph*BPL (PDB 1WQW [78]) and other fragments from the PDB using the programme "O" and Schrodinger-Prime software package [80,110], with further refinement using the PDB 1WQW in Coot [111]. The cartoon was generated in pyMOL [112] and coloured blue to red, N to C-terminal. The α-helices and β-strands are numbered as they appear in the protein sequence. The reaction intermediate biotinyl-5′-AMP is shown in the active site (in space filled representation).

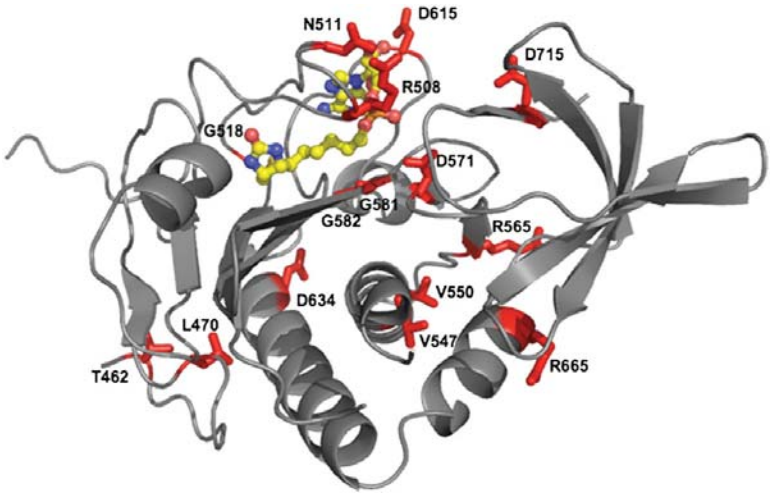


Fig. 4. Mutations leading to Multiple Carboxylase Deficiency. Residues that are mutated in MCD have been highlighted (in red) on the HCS model with biotinyl-5′-AMP shown in the active site (in stick and ball representation). It can be seen from our model that these mutations cluster around the biotin and ATP binding sites.

These patients died within the first five days of life (it is not clear whether biotin therapy was administered). Sakamoto et al. [92] reported a patient homozygous for the Leu237→Pro who has remained asymptomatic on 40 mg biotin/day. These cases suggest that the Leu237→Pro mutant does retain some small amount of biotinylation activity. The proposed mechanism by which this mutation shows some form of biotin responsiveness is due to the actual biotin concentration in the cell (estimated at 1–10 nM) being far below the reported K_M for biotin. If the enzyme possesses any residual activity, it should be biotin-responsive at higher biotin concentrations, resulting in enough enzyme activity to alleviate symptoms [85]. This suggests that V_{MAX} is an important factor in the pathology of the disease, and indeed there is a correlation between V_{MAX} and responsiveness to biotin therapy [51].

The Leu216→Arg mutation was first reported in the heterozygous form (Leu216→Arg /Val363→Asp) in [93] in a patient who showed good clinical response to biotin therapy (10–40 mg/day) [85]. Biotin responsiveness was attributed to the presence of the biotin-responsive allele Val363→Asp. Patients homozygous for the Leu216→Arg mutation have since been reported and, interestingly, their response to biotin therapy has varied [94]. This allele appears to occur at a high frequency in the Samoan and Cook Island populations. Five of the seven homozygous patients died between age 3 days to 3 years. All patients presented within 24 h of birth with severe acidosis. The two babies that did not receive biotin treatment died within seven days. Despite biotin therapy, three of the patients continued to show severe dermatological symptoms and recurrent septicaemia followed by metabolic decompensation. All died during one of these episodes. These babies also showed a lower mean birth weight (mean 2758 g vs 3700 g), indicating intrauterine growth retardation not reported for other cases of HCS deficiency. Unfortunately, extensive database searches using the N-terminal region of HCS have failed to provide evidence for homologous sequences or structures. However, a recent study demonstrated that the mutation induced increased protein turnover thereby providing a partial explanation for the poor response to biotin therapy [95]. Further studies will be required to fully understand how mutations distal from the active site of HCS have an effect on function. Throughout this review we have highlighted the importance and value of structures and the ability to use e-biology to understand more thoroughly enzymes we still know little about.

11. Further functions for HCS

In addition to the biotinylation of carboxylases, other roles have been reported for HCS including biotin-induced regulation of gene

Table 2
Structural analysis of MCD missense mutations

Mutation in HCS	Effect on structure of HCS
Y456C, T462I and L470S	Reside in an unstructured loop distal to the ligand-binding site. Importance in activity unknown.
R508W and N511K	Critical residues in loop covering the ligand-binding site. R508 co-ordinates to the backbone carbonyl N712 to form a salt bridge, removing this would result in a more flexible loop. Analogous to *Ec*BPL R118 and R121 that co-ordinate oxygens within the AMP phosphate group.
G518E	Close to active site and part of the ligand-binding loop. Mutation may not allow as much flexibility of this loop. As part of the hydrophobic pocket, making this polar residue would perturb biotin binding.
V547G and V550M	These are buried hydrophobic residues in α2 that reside near the AMP-binding β5 strand.
D571N	Important in positioning K579 in the AMP binding site.
G581S and G582R	Make hydrophobic interactions with biotin in the binding pocket.
D615Y	Reside on the loop between α3 and β6 that cover AMP. May coordinate oxygens of the AMP phosphate.
D634Y	Solvent exposed, distal to active site on α4 and α5 respectively.
D715G	On β9, may be involved in capping and stabilising the catalytic domain structure.

Using structural information about *Ec*BPL and *Ph*BPL, as well as sequence alignments with HCS, amino acids involved in MCD were identified on the HCS model. The predicted involvement of each mutation in the BPL structure and enzyme activity is stated.

expression [96]. Pioneering studies in yeast [97,98]) and rats [82] first identified this control at the transcriptional level. It was observed that biotin could influence the activity of the metabolic enzymes glucokinase and phosphoenolpyruvate carboxykinase in the livers of starved rats administered with biotin [99] by altering mRNA levels [100]. Since these early studies biotin has been shown to influence the expression of a number of genes. It has been estimated from microarray studies that as many as 1803 genes are regulated in HepG2 cells [101]. Whilst the precise biotin-sensing and transcription control mechanisms are yet to be fully defined there has been some recent activity in this field. Using small molecule inhibitors, a pathway involving a soluble guanylate cyclase and cGMP-dependent protein kinase G has been proposed to control biotin-inducible transcription, especially key enzymes in the biotin cycle such as HCS, acetyl CoA carboxylase 1, propionyl CoA carboxylase [102] and the biotin transporter SMVT [103]. Furthermore, HCS itself has been implicated in this pathway. Using fibroblasts from MCD patients homozygous for HCS Arg508→Gly (ie K_M for biotin elevated 370 fold over wildtype HCS) it was observed that 100 times more biotin was required elicit an equivalent response as in control fibroblasts [102]. The authors of this study propose that that it is biotinyl-5′-AMP, not biotin *per se*, that is a key molecule, thereby drawing parallels with the biotin-sensing pathways in yeast [104] and *E. coli* (reviewed [62]). Interestingly, the brain appears exempt from the effects of biotin starvation observed in other tissues [105]. In a rat model, it was demonstrated that biotin starvation did not affect transcription of SMVT, HCS or pyruvate carboxylase. Also, holocarboxylase levels were not changed. This was in direct contrast to liver and kidney cells where starvation correlated with decreased expression and biotinylation of carboxylases. Thus, it appears the brain has acquired specific mechanisms to ensure a continued vitamin supply during biotin stress.

Whilst the fine molecular details required to explain biotin-induced transcription remains to be uncovered, it has been proposed that HCS again plays a role. Several studies have investigated histone biotinylation, and subsequent chromatin remodelling, as one mechanism to regulate global gene expression (reviewed [106]). However, a recent report investigating techniques used for the detection of this modification to histones from cultured cells demonstrated that histone biotinylation is a very rare modification [107]. A recent study has provided an alternative mechanism using yeast as a model system. The vitamin H transporter can function as a transcription factor to regulate expression of biotin-inducible genes [108]. No mammalian homologue of the transcription factor has yet been identified but it is tempting to speculate that this story is far from complete.

References

[1] F. Kogl, B. Tonnis, Z. physiol. Chem. 242 (1936) 43.
[2] S.J. Wakil, J.K. Stoops, V.C. Joshi, Fatty acid synthesis and its regulation, Ann. Rev. Biochem. 52 (1983) 537–579.
[3] J. Zempleni, Uptake, localization, and noncarboxylase roles of biotin, Annu. Rev. Nutr. 25 (2005) 175–196.
[4] R. Spector, C.E. Johanson, Vitamin transport and homeostasis in mammalian brain: focus on Vitamins B and E, J. Neurochem. 103 (2007) 425–438.
[5] H.M. Said, Z.M. Mohammed, Intestinal absorption of water-soluble vitamins: an update, Curr. Opin. Gastroenterol. 22 (2006) 140–146.
[6] P.D. Prasad, H. Wang, R. Kekuda, T. Fujita, Y.J. Fei, L.D. Devoe, F.H. Leibach, V. Ganapathy, Cloning and functional expression of a cDNA encoding a mammalian sodium-dependent vitamin transporter mediating the uptake of pantothenate, biotin, and lipoate, J. Biol. Chem. 273 (1998) 7501–7506.
[7] H. Wang, W. Huang, Y.J. Fei, H. Xia, T.L. Yang-Feng, F.H. Leibach, L.D. Devoe, V. Ganapathy, P.D. Prasad, Human placental Na+-dependent multivitamin transporter. Cloning, functional expression, gene structure, and chromosomal localization, J. Biol. Chem. 274 (1999) 14875–14883.
[8] R.L. Daberkow, B.R. White, R.A. Cederberg, J.B. Griffin, J. Zempleni, Monocarboxylate transporter 1 mediates biotin uptake in human peripheral blood mononuclear cells, J. Nutr. 133 (2003) 2703–2706.
[9] P.T. Ozand, G.G. Gascon, M. Al Essa, S. Joshi, E. Al Jishi, S. Bakheet, J. Al Watban, M.Z. Al-Kawi, O. Dabbagh, Biotin-responsive basal ganglia disease: a novel entity, Brain 121 (Pt 7) (1998) 1267–1279.
[10] H.G. Wood, R.E. Barden, Biotin enzymes, Ann. Rev. Biochem. 46 (1977) 385–413.
[11] J.R. Knowles, The mechanism of biotin-dependent enzymes, Ann. Rev. Biochem. 58 (1989).
[12] D. Samols, C.G. Thornton, V.L. Murtif, G.K. Kumar, F.C. Haase, H.G. Wood, Evolutionary conservation among biotin enzymes, J. Biol. Chem. 263 (1988) 6461–6464.
[13] P.V. Attwood, J.C. Wallace, Chemical and catalytic mechanisms of carboxyl transfer reactions in biotin-dependent enzymes, Acc. Chem. Res. 35 (2002) 113–120.
[14] B. Wolf, Disorders of biotin metabolism, in: C.R. Scriver, A.L. Beavder, W.S. Sly (Eds.), The Metabolic and Molecular Basis of Inherited Diseases, vol. 7, McGraw-Hill, New York, 1995, pp. 3151–3177.
[15] R.W. Brownsey, A.N. Boone, J.E. Elliott, J.E. Kulpa, W.M. Lee, Regulation of acetyl-CoA carboxylase, Biochem. Soc. Trans. 34 (2006) 223–227.
[16] K.G. Thampy, Formation of malonyl coenzyme A in rat heart. Identification and purification of an isozyme of A carboxylase from rat heart, J. Biol. Chem. 264 (1989) 17631–17634.
[17] V.A. Zammit, The malonyl-CoA-long-chain acyl-CoA axis in the maintenance of mammalian cell function, Biochem. J. 343 (Pt 3) (1999) 505–515.
[18] S. Jitrapakdee, J.C. Wallace, Structure, function and regulation of pyruvate carboxylase, Biochem. J. 340 (Pt 1) (1999) 1–16.
[19] M. St Maurice, L. Reinhardt, K.H. Surinya, P.V. Attwood, J.C. Wallace, W.W. Cleland, I. Rayment, Domain architecture of pyruvate carboxylase, a biotin-dependent multifunctional enzyme, Science 317 (2007) 1076–1079.
[20] F. Deodato, S. Boenzi, F.M. Santorelli, C. Dionisi-Vici, Methylmalonic and propionic aciduria, Am. J. Med. Genet. C Semin. Med. Genet. 142 (2006) 104–112.
[21] M.R. Baumgartner, Molecular mechanism of dominant expression in 3-methyl-crotonyl-CoA carboxylase deficiency, J. Inherit. Metab. Dis. 28 (2005) 301–309.
[22] N.I. Mock, M.I. Malik, P.J. Stumbo, W.P. Bishop, D.M. Mock, Increased urinary excretion of 3-hydroxyisovaleric acid and decreased urinary excretion of biotin are sensitive early indicators of decreased biotin status in experimental biotin deficiency, Am. J. Clin. Nutr. 65 (1997) 951–958.
[23] S. Jitrapakdee, J.C. Wallace, The biotin enzyme family, conserved structural motifs and domain rearrangement, Curr. Protein Pept. Sci. (2004) 217–229.
[24] P.V. Attwood, The structure and the mechanism of action of pyruvate carboxylase, Int. J. Biochem. Cell Biol. 27 (1995) 231–249.
[25] M.D. Lane, D.L. Young, F. Lynen, The enzymatic synthesis of holotranscarboxylase from apotranscarboxylase and (+)-Biotin. I. Purification of the apoenzyme and synthetase; characteristics of the reaction, J. Biol. Chem. 239 (1964) 2858–2864.
[26] P.J. Artymiuk, D.W. Rice, A.R. Poirrette, P. Willet, A tale of two synthetases, Nat. Struct. Biol. 1 (1994) 758–760.
[27] P.A. Reche, Lipoylating and biotinylating enzymes contain a homologous catalytic module, Protein Sci. 9 (2000) 1922–1929.
[28] J.E. Cronan Jr., Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins, J. Biol. Chem. 265 (1990) 10327–10333.
[29] A. Chapman-Smith, J.E. Cronan Jr., The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity, Trends Biochem. Sci. 24 (1999) 359–363.
[30] H.C. McAllister, M.J. Coon, Further studies on the properties of liver propionyl coenzyme A holoenzyme synthetase and the specificity of holocarboxylase formation, J. Biol. Chem. 241 (1966) 2855–2861.
[31] K.E. Reed, J.E.J. Cronan, *Escherichia coli* exports previously folded and biotinated protein domains, J. Biol. Chem. 266 (1991) 11425–11428.
[32] V.L. Murtif, D. Samols, Mutagenesis affecting the carboxyl terminus of the biotinyl subunit of transcarboxylase. Effects on biotination, J. Biol. Chem. 262 (1987) 11813–11816.
[33] S.W. Polyak, A. Chapman-Smith, T.D. Mulhern, J.E. Cronan Jr., J.C. Wallace, Mutational analysis of protein substrate presentation in the post-translational attachment of biotin to biotin domains, J. Biol. Chem. 276 (2001) 3037–3045.
[34] A. Leon-Del-Rio, D. Leclerc, B. Akerman, N. Wakamatsu, R.A. Gravel, 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 (1995) 4626–4630.
[35] J.E. Cronan Jr., J.C. Wallace, The gene encoding the biotin-apoprotein ligase of Saccharomyces cerevisiae, FEMS Microbiol. Lett. 130 (1995) 221–229.
[36] E. Nenortas, D. Beckett, Purification and characterization of intact and truncated forms of the *Escherichia coli* biotin carboxyl carrier subunit of acetyl-CoA carboxylase, J. Biol. Chem. 271 (1996) 7559–7567.
[37] F.K. Athappilly, W.A. Hendrickson, Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing, Structure 3 (1995) 1407–1419.
[38] E.L. Roberts, N.C. Shu, M.J. Howard, R.W. Broadhurst, A. Chapman-Smith, J.C. Wallace, T. Morris, J.E. Cronan, R.N. Perham, 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 (1999) 5045–5053.
[39] A. Chapman-Smith, T.W. Morris, J.C. Wallace, J.E. Cronan Jr., 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 (1999) 1449–1457.
[40] C. Bru, E. Courcelle, S. Carrere, Y. Beausse, S. Dalmar, D. Kahn, The ProDom database of protein domain families: more emphasis on 3D, Nucleic Acids Res. 33 (2005) D212–D215.
[41] B.C. Shenoy, Y. Xie, V.L. Park, G.K. Kumar, H. Beegan, H.G. Wood, D. Samols, The importance of methionine residues for the catalysis of the biotin enzyme, transcarboxylase, J. Biol. Chem. 267 (1992) 18407–18412.
[42] H. Kondo, S. Uno, Y.Y. Komizo, J. Sunamoto, Importance of methionine residues in the enzymatic carboxylation of biotin containing peptides representing the local bindyl site of *E. coli* actyl CoA carboxylase, Int. J. Pept. Protein Res. 23 (1984) 559–564.

[43] P. Reche, Y.L. Li, C. Fuller, K. Eichhorn, R.N. Perham, Selectivity of post-translational modification in biotinylated proteins: the carboxy carrier protein of the acetyl-CoA carboxylase of *Escherichia coli*, Biochem. J. 329 (Pt 3) (1998) 589–596.
[44] N.G. Wallis, R.N. Perham, Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex, J. Mol. Biol. 236 (1994).
[45] J.E. Cronan Jr., The biotinyl domain of *Escherichia coli* acetyl-CoA carboxylase. Evidence that the "thumb" structure is essential and that the domain functions as a dimer, J. Biol. Chem. 276 (2001) 37355–37364.
[46] P. Reche, R.N. Perham, Structure and selectivity in post-translational modification: attaching the biotinyl-lysine and lipoyl-lysine swinging arms in multifunctional enzymes, EMBO J. 18 (1999) 2673–2682.
[47] M.M. Jank, J.D. Sadowsky, C. Peikert, S. Berger, NMR studies on the solution structure of a deletion mutant of the transcarboxylase biotin carrier subunit, Int. J. Biol. Macromol. 30 (2002) 233–242.
[48] G. Cui, B. Nan, J. Hu, Y. Wang, C. Jin, B. Xia, Identification and solution structures of a single domain biotin/lipoyl attachment protein from *Bacillus subtilis*, J. Biol. Chem. 281 (2006) 20598–20607.
[49] E. Campeau, R.A. Gravel, 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 (2001) 12310–12316.
[50] G. Tissot, R. Douce, C. Alban, 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 (1997) 179–188.
[51] O. Sakamoto, Y. Suzuki, X. Li, Y. Aoki, M. Hiratsuka, T. Suormala, E.R. Baumgartner, K.M. Gibson, K. Narisawa, Relationship between kinetic properties of mutant enzyme and biochemical and clinical responsiveness to biotin in holocarboxylase synthetase deficiency, Pediatr. Res. 46 (1999) 671–676.
[52] S.W. Polyak, A. Chapman-Smith, P.J. Brautigan, J.C. Wallace, Biotin protein ligase from *Saccharomyces cerevisiae*. The N-terminal domain is required for complete activity, J. Biol. Chem. 274 (1999) 32847–32854.
[53] E.R. Baumgartner, T. Suormala, Multiple carboxylase deficiency—inherited and acquired disorders of biotin metabolism, Int. J. Vitam. Nutr. Res. 67 (1997) 377–384.
[54] M. Saunders, L. Sweetman, B. Robinson, K. Roth, R. Cohn, R.A. Gravel, Biotin-response organicaciduria. Multiple carboxylase defects and complementation studies with propionicacidemia in cultured fibroblasts, J. Clin. Invest. 64 (1979) 1695–1702.
[55] B.J. Burri, L. Sweetman, W.L. Nyhan, Mutant holocarboxylase synthetase: evidence for the enzyme defect in early infantile biotin-responsive multiple carboxylase deficiency, J. Clin. Invest. 68 (1981) 1491–1495.
[56] J. Thoene, B. Wolf, Biotinidase deficiency in juvenile multiple carboxylase deficiency, Lancet 2 (1983) 398.
[57] B. Wolf, R.E. Grier, J.R. Secor McVoy, G.S. Heard, Biotinidase deficiency: a novel vitamin recycling defect, J. Inherit. Metab. Dis. 8 (Suppl 1) (1985) 53–58.
[58] O. Sakamoto, Y. Suzuki, X. Li, Y. Aoki, M. Hiratsuka, E. Holme, J. Kudoh, N. Shimizu, K. Narisawa, Diagnosis and molecular analysis of an atypical case of holocarboxylase synthetase deficiency, Eur. J. Pediatr. 159 (2000) 18–22.
[59] Y. Suzuki, X. Yang, Y. Aoki, S. Kure, Y. Matsubara, Mutations in the holocarboxylase synthetase gene HLCS, Human Mutat. 26 (2005) 285–290.
[60] A.M. Lund, F. Joensen, D.M. Hougaard, L.K. Jensen, E. Christensen, M. Christensen, B. Norgaard-Petersen, M. Schwartz, F. Skovby, Carnitine transporter and holocarboxylase synthetase deficiencies in The Faroe Islands, J. Inherit. Metab. Dis. 30 (2007) 341–349.
[61] S. Malvagia, A. Morrone, E. Pasquini, S. Funghini, G. la Marca, E. Zammarchi, M.A. Donati, First prenatal molecular diagnosis in a family with holocarboxylase synthetase deficiency, Prenat. Diagn. 25 (2005) 1117–1119.
[62] D. Beckett, Biotin sensing: universal influence of biotin status on transcription, Annu. Rev. Genet. 41 (2007) 443–464.
[63] K.P. Wilson, L.M. Shewchuk, R.G. Brennan, A.J. Otsuka, B.W. Matthews, *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 9257–9261.
[64] A. Chapman-Smith, T.D. Mulhern, F. Whelan, J.E. Cronan Jr., J.C. Wallace, The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity, Protein Sci. 10 (2001) 2608–2617.
[65] L.H. Weaver, K. Kwon, D. Beckett, B.W. Matthews, Competing protein:protein interactions are proposed to control the biological switch of the *E. coli* biotin repressor, Protein Sci. 10 (2001) 2618–2622.
[66] L.H. Weaver, K. Kwon, D. Beckett, B.W. Matthews, 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 (2001) 6045–6050.
[67] A. Morrone, S. Malvagia, M.A. Donati, S. Funghini, F. Ciani, I. Pela, A. Boneh, H. Peters, E. Pasquini, E. Zammarchi, Clinical findings and biochemical and molecular analysis of four patients with holocarboxylase synthetase deficiency, Am. J. Med. Genet. 111 (2002) 10–18.
[68] C.R. Bellamacina, The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins, FASEB J. 10 (1996) 1257–1269.
[69] K. Kwon, D. Beckett, Function of a conserved sequence motif in biotin holoenzyme synthetases, Protein Sci. 9 (2000) 1530–1539.
[70] Z.A. Wood, L.H. Weaver, P.H. Brown, D. Beckett, B.W. Matthews, Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution, J. Mol. Biol. 357 (2006) 509–523.
[71] S. Naganathan, D. Beckett, Nucleation of an allosteric response via ligand-induced loop folding, J. Mol. Biol. 373 (2007) 96–111.
[72] Y. Xu, C.R. Johnson, D. Beckett, Thermodynamic analysis of small ligand binding to the *Escherichia coli* repressor of biotin biosynthesis, Biochemistry 35 (1996) 5509–5517.
[73] A.J. Otsuka, *The Escherichia coli* biotin biosynthetic enzyme sequence predicted from the nucleotide sequence of the bio operon, J. Biol. Chem. (1988) 19577–19585.
[74] M.A. Eisenberg, Biotin: biogenesis, transport and their regulation, Adv. Enzymol. 38 (1973) 317–372.
[75] J. Abbott, D. Beckett, Cooperative binding of the *Escherichia coli* repressor of biotin biosynthesis to the biotin operator sequence, Biochemistry 32 (1993) 9649–9656.
[76] E. Eisenstein, D. Beckett, Dimerization of the *Escherichia coli* biotin repressor: corepressor function in protein assembly, Biochemistry 38 (1999) 13077–13084.
[77] A.J. Otsuka, J. Abelson, The regulatory region of the biotin operon in *Escherichia coli*, Nature (1978) 689–693.
[78] B. Bagautdinov, C. Kuroishi, M. Sugahara, N. Kunishima, Crystal structures of biotin protein ligase from *Pyrococcus horikoshii* OT3 and its complexes: structural basis of biotin activation, J. Mol. Biol. 353 (2005) 322–333.
[79] B. Bagautdinov, Y. Matsuura, S. Bagautdinova, N. Kunishima, Crystallization and preliminary X-ray crystallographic studies of the biotin carboxyl carrier protein and biotin protein ligase complex from *Pyrococcus horikoshii* OT3, Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 63 (2007) 334–337.
[80] L. Schrodinger, Prime, Schrodinger, L.L.C, Portland (Oregon), 2005.
[81] G.J.K.T.A. Jones, Databases in protein crystallography, Acta Crystallogr. D34 (1998) 1119–1131.
[82] B.J. Burri, L. Sweetman, W.L. Nyhan, Heterogeneity of holocarboxylase synthetase in patients with biotin-responsive multiple carboxylase deficiency, Am. J. Hum. Genet. 37 (1985) 326–337.
[83] J. Morita, L.P. Thuy, L. Sweetman, Deficiency of biotinyl-AMP synthetase activity in fibroblasts of patients with holocarboxylase synthetase deficiency, Mol. Genet. Metab. 64 (1998) 250–255.
[84] Y. Suzuki, Y. Aoki, O. Sakamoto, X. Li, S. Miyabayashi, Y. Kazuta, H. Kondo, K. Narisawa, Enzymatic diagnosis of holocarboxylase synthetase deficiency using apo-carboxyl carrier protein as a substrate, Clin. Chim. Acta 251 (1996) 41–52.
[85] L. Dupuis, E. Campeau, D. Leclerc, R.A. Gravel, Mechanism of biotin responsiveness in biotin-responsive multiple carboxylase deficiency, Mol. Genet. Metab. 66 (1999) 80–90.
[86] K. Kwon, E.D. Streaker, D. Beckett, Binding specificity and the ligand dissociation process in the *E. coli* biotin holoenzyme synthetase, Protein Sci. 11 (2002) 558–570.
[87] E. Choi-Rhee, H. Schulman, J.E. Cronan, Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase, Protein Sci. 13 (2004) 3043–3050.
[88] Y. Aoki, X. Li, O. Sakamoto, M. Hiratsuka, H. Akaishi, L. Xu, P. Briones, T. Suormala, E.R. Baumgartner, Y. Suzuki, K. Narisawa, Identification and characterization of mutations in patients with holocarboxylase synthetase deficiency, Hum. Genet. 104 (1999) 143–148.
[89] Y. Aoki, Y. Suzuki, O. Sakamoto, X. Li, K. Takahashi, A. Ohtake, R. Sakuta, T. Ohura, S. Miyabayashi, K. Narisawa, Molecular analysis of holocarboxylase synthetase deficiency — a missense mutation and a single base deletion are predominant In Japanese patients, Biochim. Biophys. Acta, Mol. Basis Dis. 1272 (1995) 168–174.
[90] Y. Aoki, Y. Suzuki, X. Li, O. Sakamoto, H. Chikaoka, S. Takita, K. Narisawa, Characterization of mutant holocarboxylase synthetase (HCS)—a *K*m for biotin was not elevated in a patient with HCS deficiency, Pediatr. Res. 42 (1997) 849–854.
[91] X. Yang, Y. Aoki, X. Li, O. Sakamoto, M. Hiratsuka, K.M. Gibson, S. Kure, K. Narisawa, Y. Matsubara, Y. Suzuki, Haplotype analysis suggests that the two predominant mutations in Japanese patients with holocarboxylase synthetase deficiency are founder mutations, J. Hum. Genet. 45 (2000) 358–362.
[92] O. Sakamoto, Y. Suzuki, Y. Aoki, X. Li, M. Hiratsuka, K. Yanagihara, K. Inui, T. Okabe, S. Yamaguchi, J. Kudoh, N. Shimizu, K. Narisawa, Molecular analysis of new Japanese patients with holocarboxylase synthetase deficiency, J. Inherit. Metab. Dis. 21 (1998) 873–874.
[93] L. Dupuis, A. Leon-Del-Rio, D. Leclerc, E. Campeau, L. Sweetman, J.M. Saudubray, G. Herman, K.M. Gibson, R.A. Gravel, Clustering of mutations in the biotin-binding region of holocarboxylase synthetase in biotin-responsive multiple carboxylase deficiency, Hum. Mol. Genet. 5 (1996) 1011–1016.
[94] C.J. Wilson, M. Myer, B.A. Darlow, T. Stanley, G. Thomson, E.R. Baumgartner, D.M. Kirby, D.R. Thorburn, Severe holocarboxylase synthetase deficiency with incomplete biotin responsiveness resulting in antenatal insult in Samoan neonates, J. Pediatr. 147 (2005) 115–118.
[95] L.M. Bailey, R.A. Ivanov, S. Jitrapakdee, C.J. Wilson, J.C. Wallace, Reduced half-life of holocarboxylase synthetase from patients with severe multiple carboxylase deficiency, Human Mutat. 1009 (2) (2008) E47–E57 (Mutation in Brief).
[96] D. Pacheco-Alvarez, R.S. Solorzano-Vargas, A.L. Del Rio, Biotin in metabolism and its relationship to human disease, Arch. Med. Res. 33 (2002) 439–447.
[97] J. Stolz, Isolation and characterization of the plasma membrane biotin transporter from *Schizosaccharomyces pombe*, Yeast 20 (2003) 221–231.
[98] J. Stolz, U. Hoja, S. Meier, N. Sauer, E. Schweizer, Identification of the plasma membrane H+-biotin symporter of *Saccharomyces cerevisiae* by rescue of a fatty acid-auxotrophic mutant, J. Biol. Chem. 274 (1999) 18741–18746.
[99] K. Dakshinamurti, P.R. Desjardins, Acetyl-CoA carboxylase from rat adipose tissue, Biochim. Biophys. Acta 176 (1969) 221–229.
[100] J. Chauhan, K. Dakshinamurti, Transcriptional regulation of the glucokinase gene by biotin in starved rats, J. Biol. Chem. 266 (1991) 10035–10038.
[101] R. Rodriguez-Melendez, J.B. Griffin, G. Sarath, J. Zempleni, High-throughput immunoblotting identifies biotin-dependent signaling proteins in HepG2 hepatocarcinoma cells, J. Nutr. 135 (2005) 1659–1666.
[102] R.S. Solorzano-Vargas, D. Pacheco-Alvarez, A. Leon-Del-Rio, Holocarboxylase synthetase is an obligate participant in biotin-mediated regulation of its own

expression and of biotin-dependent carboxylases mRNA levels in human cells, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 5325–5330.
[103] D. Pacheco-Alvarez, R.S. Solorzano-Vargas, A. Gonzalez-Noriega, C. Michalak, J. Zempleni, A. Leon-Del-Rio, Biotin availability regulates expression of the sodium-dependent multivitamin transporter and the rate of biotin uptake in HepG2 cells, Mol. Genet. Metab. 85 (2005) 301–307.
[104] H.M. Pirner, J. Stolz, Biotin sensing in *Saccharomyces cerevisiae* is mediated by a conserved DNA element and requires the activity of biotin-protein ligase, J. Biol. Chem. 281 (2006) 12381–12389.
[105] D. Pacheco-Alvarez, R.S. Solorzano-Vargas, R.A. Gravel, R. Cervantes-Roldan, A. Velazquez, A. Leon-Del-Rio, Paradoxical regulation of biotin utilization in brain and liver and implications for inherited multiple carboxylase deficiency, J. Biol. Chem. 279 (2004) 52312–52318.
[106] N. Kothapalli, G. Camporeale, A. Kueh, Y.C. Chew, A.M. Oommen, J.B. Griffin, J. Zempleni, Biological functions of biotinylated histones, J. Nutr. Biochem 16 (2005) 446–448.
[107] L.M. Bailey, R.A. Ivanov, J.C. Wallace, S.W. Polyak, Artifactual detection of biotin on histones by streptavidin, Anal. Biochem. 373 (2008) 71–77.
[108] M. Weider, A. Machnik, F. Klebl, N. Sauer, Vhr1p, a new transcription factor from budding yeast, regulates biotin-dependent expression of VHT1 and BIO5, J. Biol. Chem. 281 (2006) 13513–13524.
[109] D.V. Reddy, B.C. Shenoy, P.R. Carey, F.D. Sonnichsen, High resolution solution structure of the 1.3S subunit of transcarboxylase from *Propionibacterium shermanii*, Biochemistry 39 (2000) 2509–2516.
[110] T.A. Jones, J.Y. Zou, S.W. Cowan, M. Kjeldgaard, Improved methods for building protein models in electron density maps and the location of errors in these models, Acta Crystallogr., A 47 (Pt 2) (1991) 110–119.
[111] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr., D Biol. Crystallogr. 60 (2004) 2126–2132.
[112] W. DeLano, The PyMOL molecular graphics system.
[113] X. Yang, Y. Aoki, X. Li, O. Sakamoto, M. Hiratsuka, S. Kure, S. Taheri, E. Christensen, K. Inui, M. Kubota, M. Ohira, M. Ohki, J. Kudoh, K. Kawasaki, K. Shibuya, A. Shintani, S. Asakawa, S. Minoshima, N. Shimizu, K. Narisawa, Y. Matsubara, Y. Suzuki, Structure of human holocarboxylase synthetase gene and mutation spectrum of holocarboxylase synthetase deficiency, Hum. Genet. 109 (2001) 526–534.

Chapter 3

Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from *Staphylococcus aureus*.

This chapter describes the cloning, expression and purification of the Biotin Protein Ligase from *Staphylococcus aureus* (SaBPL). This allowed for crystallisation trials to be conducted. Fortunately, protein crystals were produced of the enzyme and data collected and analysed. Methods and programmes used for each step to obtain the data required for solving the structure of SaBPL are detailed here. A collaboration between the University of Adelaide and Monash University was established to achieve these works due to lack of equipment and expertise available in Adelaide.

Chapter 3

Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from *Staphylococcus aureus*.

Author	Contribution	Signature
Nicole R. Pendini	Primary role in performing experiments and manuscript preparation. Performed protein expression, purification and crystallisation screens, optimisation and data collection. Produced figures 2 and 3 and Table 1.	
*Steven W. Polyak	Performed all recombinant DNA cloning including construction of expression vectors. Optimised the protocol for protein purification. Provided Figure 1 and contributed to manuscript preparation.	
*Grant W. Booker	Provided intellectual discussion and assisted in manuscript preparation.	
*John C. Wallace	Provided intellectual discussion and assisted in manuscript preparation.	
Matthew C. Wilce	Provided direction and assistance in all aspects of X-ray crystallography, data collection, analysis and manuscript preparation.	

This work was supported through a Commercial Development Initiative grant from BioInnovationSA awarded to SWP, GWB and JCW.

Permission for reprint from Acta Crystallographica section F

Dear Nicole

Permission is hereby granted, on behalf of the IUCr, for you to reproduce the article Acta Cryst. (2008). F64, 520-523 [doi:10.1107/S1744309108012244] Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from *Staphylococcus aureus* by N. R. Pendini, S. W. Polyak, G. W. Booker, J. C. Wallace and M. C. J. Wilce in your PhD thesis.

This permission is subject to the following conditions:

* Reproduction is intended in a primary journal, secondary journal, CD-ROM, book or thesis. * The original article is cited.

* IUCr's copyright permission is indicated by the wording "Reproduced with permission of the International Union of Crystallography". In electronic form, this acknowledgement must be visible at the same time as the reused materials, and must be hyperlinked to Crystallography Journals Online (http://journals.iucr.org/).

Best wishes

Simon Glynn (Technical Editor)

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

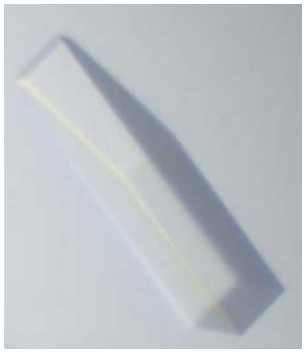
Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from *Staphylococcus aureus*

Nicole R. Pendini,[a,b] Steve W. Polyak,[a] Grant W. Booker,[a] John C. Wallace[a] and Matthew C. J. Wilce[b]*

[a]School of Molecular and Biomedical Sciences, University of Adelaide, North Terrace, Adelaide SA 5005, Australia, and [b]Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton VIC 3800, Australia

Correspondence e-mail: matthew.wilce@med.monash.edu.au

Received 8 April 2008
Accepted 28 April 2008



Biotin protein ligase from *Staphylococcus aureus* catalyses the biotinylation of acetyl-CoA carboxylase and pyruvate carboxylase. Recombinant biotin protein ligase from *S. aureus* has been cloned, expressed and purified. Crystals were grown using the hanging-drop vapour-diffusion method using PEG 8000 as the precipitant at 295 K. X-ray diffraction data were collected to 2.3 Å resolution from crystals using synchrotron X-ray radiation at 100 K. The diffraction was consistent with the tetragonal space group $P4_22_12$, with unit-cell parameters $a = b = 93.665$, $c = 131.95$.

1. Introduction

Staphylococcus aureus is a clinically important opportunistic pathogen that has acquired resistance to almost all antibiotics on the market, including penicillin, methicillin and vancomycin (Gomes *et al.*, 2006). There is a desperate need to develop new therapeutic agents to target this pathogen. These agents may be a novel class of compounds that target differing pathways to those already on the market. Biotin protein ligase (BPL) is one such target (Payne *et al.*, 2007) as it is an essential enzyme (Cronan & Wallace, 1995) that is responsible for the post-translational attachment of biotin to biotin-dependent enzymes. Biotinylation occurs in a two-step reaction in which BPL produces the reaction intermediate biotinyl-5′-AMP from biotin and ATP before transferring the biotin moiety onto a specific lysine residue in the active site of the biotin-dependent enzyme (Lane *et al.*, 1964). *S. aureus* expresses two biotin-requiring enzymes, namely acetyl-CoA carboxylase and pyruvate carboxylase, that participate in important metabolic pathways. Acetyl-CoA carboxylase catalyses the production of malonyl-CoA from acetyl-CoA, which is the first committed step in fatty-acid biosynthesis (Wakil *et al.*, 1983). Pyruvate carboxylase is a gluconeogenic enzyme that catalyses the conversion of pyruvate to oxaloacetate (Wallace *et al.*, 1998; Attwood & Keech, 1984). The structure of *S. aureus* pyruvate carboxylase has recently been determined (Xiang & Tong, 2008). Whilst the biotin enzymes have been identified as targets for new antibiotics (Forsyth *et al.*, 2002), exploiting BPL in the same manner has yet to be reported.

To date, structures of BPL from the Gram-negative bacterium *Escherichia coli* (EcBPL; Wood *et al.*, 2006; Weaver *et al.*, 2001; Wilson *et al.*, 1992) and the archeon *Pyrococcus horikoshii* (PhBPL; Bagautdinov *et al.*, 2005, 2008) have been well characterized. The coordinates for the BPL structures from *Mycobacterium tuberculosis* (PDB code 2cgh), *Aquifex aeolicus* (PDB code 2eay) and *Methanococcus jannaschii* (PDB code 2ej9) have been deposited in the PDB, but no analyses of these structures have been published. All BPL structures solved thus far have a conserved catalytic core composed of two domains with structural homology to SH2 and SH3 domains. Residues within the SH2-like domain are required for binding biotin, whereas the C-terminal SH3-like domain is required for nucleotide binding and forms a cap over the active site upon ligand binding

(Wilson *et al.*, 1992; Beckett, 2005; Chapman-Smith *et al.*, 2001). The microbial BPLs can be divided into two separate classes. The class I BPLs are simple two-domain structures considered to be the minimal catalytic unit. The BPLs from *P. horikoshii*, *My. tuberculosis*, *A. aeolicus* and *Me. jannaschii* all belong to this class. Interestingly, PhBPL is a constitutive homodimer, although the biological significance of this is not understood. The class II BPLs are more complex and contain an additional N-terminal domain that is required for DNA binding. EcBPL, commonly known as the biotin-inducible repressor protein (BirA), has been extensively investigated by genetic, mutational and structural studies. BirA is monomeric in the absence of ligand, but can be induced to form a homodimer upon binding of biotin or biotinyl-5′-AMP. The orientation of the N-terminal domains in the homodimer facilitates binding to a specific DNA sequence that is present in the operator region of the biotin-biosynthetic operon, where it functions as a transcriptional repressor. Thus, BirA is a bifunctional protein that regulates both the synthesis and utilization of biotin. Importantly, the PhBPL and BirA dimers are structurally distinct and employ different surfaces for dimerization. From DNA-sequence analysis, the BPL from *S. aureus* (SaBPL) is proposed to contain an N-terminal DNA-binding domain and to belong to class II. However, SaBPL has a higher sequence identity to PhBPL (29% compared with 22% for EcBPL). Here, we present the expression, purification and crystallization of a BPL from a Gram-positive bacterium. Detailed knowledge of the SaBPL structure may facilitate the discovery and development of new therapeutic agents.

2. Methods and results

2.1. DNA manipulations

The gene for *S. aureus* BPL was obtained by genomic PCR using an isolate of *S. aureus*. Oligonucleotides were designed using the DNA sequence deposited in GenBank (accession No. NP_371980; gi:15924446; genomic sequence of methicillin-resistant *S. aureus*; Kuroda *et al.*, 2001). To facilitate subsequent cloning, a *Pci*I restriction site was engineered into the 5′ primer (B138/38, 5′-ACATGTCAAAATATAGTCAAGATGTACTTCAATTACTC) and *Bam*HI and *Hin*dIII restriction sites were engineered into the 3′ primer (B139/40, 5′-GGATCCAAGCTTAAAAATCTATATCTGCACTAATAAAACG). PCR was performed under standard conditions (Ling *et al.*, 1991) using either Vent DNA polymerase (New England Biolabs) for genomic PCR or Dynazyme EXT DNA polymerase (New England Biolabs) to link single adenosine bases onto the 5′ ends of the fragment for TA cloning. Plasmid pGEM(*S. aureus* BPL) was generated by ligating the resulting 1 kbp fragment from genomic PCR into pGEM-T Easy vector (Promega). To facilitate the purification of SaBPL by metal-chelating chromatography, a hexahistidine sequence was engineered onto the C-terminus of the protein. This was performed using PCR with oligonucleotides B138/38 and B140/55 (5′-AAGCTTAATGATGATGATGATGACCAAAATCTATATCTGCACTAATTAAACG) and pGEM(*S. aureus* BPL) as a template. The PCR fragment was cloned into pGEM-T Easy, yielding pGEM(*S. aureus* BPL-H_6). For overexpression of recombinant *S. aureus* BPL-H_6 in *E. coli*, the 1 kbp fragment liberated from pGEM(*S. aureus* BPL-H_6) upon digestion with *Pci*I and *Hin*dIII was ligated into *Nco*I- and *Hin*dIII-treated pET16b (Novagen). The resulting plasmid, pET(*S. aureus* BPL-H_6), was transformed into *E. coli* BL21 (DE3) for recombinant protein expression. All constructs were confirmed by DNA sequencing. The isolated DNA sequence was confirmed to be a functional BPL using a complementation assay with the temperature-sensitive *birA*$^-$ mutant *E. coli* strain CY218 (Chapman-Smith *et al.*, 1994).

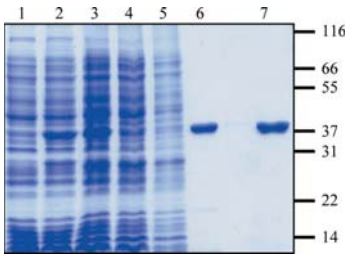


Figure 1
Purification of recombinant SaBPL-H_6. The purification strategy was monitored by SDS–PAGE on a 12% polyacrylamide gel. The His_6-tagged protein migrated at a relative mobility of ∼38 kDa. Lane 1 contains crude cell lysate before recombinant protein expression, lane 2 contains crude lysate after induction with 1 m*M* IPTG and lane 3 contains 50 µg of the soluble fraction after cell lysis. SaBPL-H_6 was then purified by IMAC chromatography (lane 4, 50 µg unbound protein; lane 5, 50 µg of the wash fraction; lane 6, 5 µg eluted protein) and S-Sepharose ion-exchange chromoatography (lane 7, 5 µg of SaBPL-H_6 from pooled fractions containing BPL activity). Molecular-weight standards (labelled in kDa) are shown on the right.

2.2. Recombinant protein expression and purification

E. coli BL21 (DE3) cells harbouring pET(*S. aureus* BPL-H_6) were grown at 310 K in 2YT medium supplemented with 200 µg ml^{-1} ampicillin to an OD_{600} of 0.6. Recombinant protein expression was induced with 1 m*M* IPTG for 4 h at 303 K. The cells were harvested by centrifugation at 2968*g* for 20 min at 277 K before resuspension in buffer *A* (20 m*M* Tris–HCl, 0.5 *M* NaCl pH 7.9, 20 m*M* imidazole) with 1 m*M* PMSF. Cells were lysed by sonication and French press. Cellular debris was removed by centrifugation at 2968*g* for 20 min. SaBPL-H_6 was then purified using immobilized metal-affinity chromatography (IMAC). The filtered supernatant was applied onto a 5 ml His-Trap HP column (GE Healthcare) equilibrated with buffer *A*. The column was washed with ten column volumes of buffer *A* followed by three column volumes of buffer *A* containing 50 m*M* imidazole before SaBPL-H_6 was eluted with buffer *A* containing 100 m*M* imidazole. Fractions containing SaBPL-H_6 were determined by SDS–PAGE and Western blotting probed with Ni–NTA alkaline phosphatase. These fractions were pooled, concentrated and buffer-exchanged into buffer *B* (50 m*M* sodium phosphate, 5% glycerol, 1 m*M* EDTA and 1 m*M* DTT) using an Amicon Ultra Centrifugal Filter Device (Millipore).

Analysis of the IMAC-purified protein by SDS–PAGE showed that the sample was not homogeneous. Ion-exchange chromatography has previously been employed for purification of *E. coli* and yeast BPL (Polyak *et al.*, 1999; Chapman-Smith *et al.*, 2001). Therefore, the IMAC-purified protein was applied onto a 15 ml SP Sepharose Fast Flow (GE Healthcare) column equilibrated in buffer *B*. The sample was then fractionated using a 0–400 m*M* NaCl gradient over 40 min, with SaBPL-H_6 eluting at 200 m*M* NaCl. Fractions containing SaBPL-H_6 were confirmed using an *in vitro* biotinylation assay (Chapman-Smith *et al.*, 1999). SDS–PAGE revealed the protein to be homogenous (Fig. 1). Mass spectrometry measured the protein mass as 37 891.3 (the predicted M_r was 37 893.4). Following concentration, SaBPL-H_6 was stored at 1–2 mg ml^{-1} in storage buffer (50 m*M* Tris–HCl pH 7.5, 0.5 m*M* EDTA pH 8.0, 250 m*M* NaCl) at 193 K. The enzyme retained catalytic activity for at least six months.

2.3. Crystallization

SaBPL-H_6 was concentrated to 12.5 mg ml^{-1} using a Vivaspin 10 kDa molecular-weight cutoff concentrator (Sartorius Stedim Biotech, Victoria, Australia) and buffer-exchanged into 20 m*M* Tris–HCl pH 7.5. Initial crystallization conditions were found using the Sigma protein–protein interaction screen (Sigma–Aldrich, NSW, Australia). In brief, 500 µl of each reagent from the 48 pre-composed Sigma protein–protein interaction screen conditions were dispensed into Limbro tissue-culture plates. 1 µl protein-solution drops were mixed with 1 µl reservoir solution and equilibrated over the reservoir at 293 K on Hampton siliconized 22 mm glass slides using the hanging-drop vapour-diffusion method. Crystals were observed in two conditions: 20% PEG 3350, 0.1 *M* MOPS pH 7.5, 0.1 *M* MgCl and 12% PEG 8000, 0.1 *M* Tris–HCl pH 8.0 (condition Nos. 11 and 34 respectively). The first condition produced thin plate-shaped crystals (0.05 × 0.05 × 0.01 mm) that bundled together, while the second condition produced rectangular brick-like single crystals (approximate dimensions 0.2 × 0.15 × 0.1 mm). Both crystal forms gave rise to X-ray diffraction that was consistent with protein crystals. Crystallization optimization was performed by screening PEG molecular weights (2000–120 000) and concentrations (6–22%), the pH of the buffer (7–9.5) and the type of buffer (including HEPES, MES and MOPS) used for each condition and varying the protein concentration (10–40 mg ml^{-1}). The optimum crystals grew using a 500 µl reservoir containing 8% PEG 8000 and 0.1 *M* Tris–HCl pH 8.0. 1 µl protein solution (15 mg ml^{-1}) was mixed with 1 µl reservoir solution and equilibrated at either 277 or 293 K using the hanging-drop vapour-diffusion method. Crystals appeared within 24 h at both temperatures, with dimensions of approximately 0.3 × 0.2 × 0.1 mm. A single crystal was picked up using a Hampton silicon loop and streaked through cryoprotectant solutions containing a 1:4, a 1:3 and then a 1:2 ratio of 100% glycerol:reservoir condition and flash-cooled at 100 K. X-ray diffraction data were collected on the high-throughput protein crystallography beamline at the Australian Synchrotron using a MAR 165 CCD detector. 180 diffraction images were recorded. The oscillation angle for each frame was 0.5° and the exposure time was 20 s. The diffraction data were integrated using *MOSFLM* (Leslie, 1999) and the intensities were merged and scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994). Wilson scaling was applied using *TRUNCATE* (Collaborative Computational Project, Number 4, 1994). Molecular replacement is under way.

Table 1
Crystallographic data and refinement statistics for SaBPL.

Values in parentheses are for the highest resolution bin.

Wavelength (Å)	0.95367
Resolution range (Å)	29–2.33 (2.33–2.4)
Total observations	543518 (15545)
Unique reflections	30148 (2076)
Completeness (%)	100 (99.6)
Redundancy	7.4 (7.1)
Mean $I/\sigma(I)$	8.9 (2.3)
Unit-cell parameters (Å, °)	a = 93.665, b = 93.665, c = 131.95, $\alpha = \beta = \gamma$ = 90
$R_{\rm p.i.m.}$† (%)	5.0 (44.9)

† $R_{\rm p.i.m.}$ is the precision-indicating (multiplicity-weighted) $R_{\rm merge}$ (Diederichs & Karplus, 1997).

3. Results and discussion

The gene for BPL from *S. aureus* was isolated and cloned into an expression system for recombinant expression in *E. coli*. Using two chromatography steps, the protein was purified to homogeneity for crystallization. The purification procedure was optimized to facilitate rapid isolation of the enzyme. This allowed the protein to be rapidly purified from cell lysates in 1 d, thus minimizing storage and freeze–thawing of the enzyme. Using this homogeneous protein, crystals of SaBPL-H_6 were grown and X-ray diffraction data were collected. Crystals appeared within 24 h, with some crystals growing as large as 0.3 × 0.2 × 0.1 mm (Fig. 2). A crystal grown at 293 K degraded within 24 h, while crystals grown at 277 K were stable for up to one week. The crystals produced at 277 K were used for diffraction data collection (Fig. 3). The diffraction data were consistent with space group *P*422. Systematic absences suggested that the crystals belonged to space group $P4_22_12$. Based upon the Matthews coefficient, $V_{\rm M}$ =

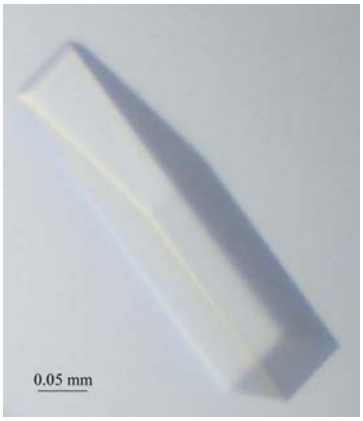


Figure 2
Crystal of biotin protein ligase from *S. aureus* of approximately 0.3 × 0.2 × 0.1 mm in size.

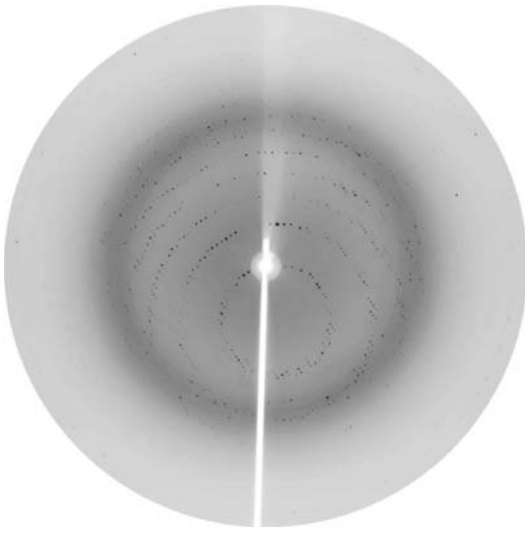


Figure 3
X-ray diffraction from a native *S. aureus* BPL crystal. The edge of the detector corresponds to 2.16 Å resolution.

2.16 Å^3 Da^{-1}, which is consistent with approximately 40% solvent content, we anticipate that there will be one molecule in the asymmetric unit. Diffraction data are summarized in Table 1.

We wish to thank Dr Renato Morona for access to the *S. aureus* bacterium for cloning and Julian Adams and the staff at the Protein Crystallography beamline at the Australian Synchrotron for their training and guidance. This work was supported by the Biochemistry Discipline, School of Molecular and Biomedical Sciences, University of Adelaide. SWP, GWB and JCW were awarded a Commercial Development Initiative grant, BioInnovationSA. MCJW acknowledges support from the National Health and Medical Research Council and the Australian Research Council.

References

Attwood, P. V. & Keech, D. B. (1984). *Curr. Top. Cell Regul.* **23**, 1–55.
Bagautdinov, B., Kuroishi, C., Sugahara, M. & Kunishima, N. (2005). *J. Mol. Biol.* **353**, 322–333.
Bagautdinov, B., Matsuura, Y., Bagautdinova, S. & Kunishima, N. (2008). *J. Biol. Chem.* doi: 10.1074/jbc.M709116200.
Beckett, D. (2005). *J. Nutr. Biochem.* **16**, 411–415.
Chapman-Smith, A., Morris, T. W., Wallace, J. C. & Cronan, J. E. Jr (1999). *J. Biol. Chem.* **274**, 1449–1457.
Chapman-Smith, A., Mulhern, T. D., Whelan, F., Cronan, J. E. Jr & Wallace, J. C. (2001). *Protein Sci.* **10**, 2608–2617.
Chapman-Smith, A., Turner, D. L., Cronan, J. E., Morris, T. W. & Wallace, J. C. (1994). *Biochem. J.* **302**, 881–887.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cronan, J. E. Jr & Wallace, J. C. (1995). *FEMS Microbiol. Lett.* **130**, 221–229.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Forsyth, R. A. *et al.* (2002). *Mol. Microbiol.* **43**, 1387–1400.
Gomes, A. R., Westh, H. & de Lencastre, H. (2006). *Antimicrob. Agents Chemother.* **50**, 3237–3244.
Kuroda, M. *et al.* (2001). *Lancet*, **357**, 1225–1240.
Lane, M. D., Rominger, K. L., Young, D. L. & Lynen, F. (1964). *J. Biol. Chem.* **239**, 2865–2871.
Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.
Ling, L. L., Keohavong, P., Dias, C. & Thilly, W. G. (1991). *PCR Methods Appl.* **1**, 63–69.
Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. (2007). *Nature Rev. Drug Discov.* **6**, 29–40.
Polyak, S. W., Chapman-Smith, A., Brautigan, P. & Wallace, J. C. (1999). *J. Biol. Chem.* **274**, 32847–32854.
Wakil, S. J., Stoops, J. K. & Joshi, V. C. (1983). *Annu. Rev. Biochem.* **52**, 537–579.
Wallace, J. C., Jitrapakdee, S. & Chapman-Smith, A. (1998). *Int. J. Biochem. Cell Biol.* **30**, 1–5.
Weaver, L. H., Kwon, K., Beckett, D. & Matthews, B. W. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 6045–6050.
Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 9257–9261.
Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D. & Matthews, B. W. (2006). *J. Mol. Biol.* **357**, 509–523.
Xiang, S. & Tong, L. (2008). *Nature Struct. Mol. Biol.* **15**, 295–302.

Chapter 4

Crystal structures of apo and liganded Biotin Protein Ligase from *Staphylococcus aureus*: towards the development of new antibiotics for MRSA.

This chapter takes an involved look into the structure of *S. aureus* BPL. I have solved the structures of the unliganded (apo), biotin bound and biotinyl-5'-AMP (holo) forms by using the method of molecular replacement. The structure has been extensively analysed both in ligand and unliganded forms. Differences and similarities between these forms have been documented in the first part of the chapter followed by an in depth analysis of the structural correlation to other BPL structures that have been deposited in the protein data bank. These analysis required many different programs and methods that have been referenced throughout the text.

Chapter 4:

Crystal structures of apo and liganded biotin protein ligase from *Staphylococcus aureus:* towards the development of new antibiotics for MRSA.

Author	Contribution	Signature
Nicole R. Pendini	Primary role in both performing experiments and writing of the manuscript. Performed crystallisation screens, optimisation, ligand screens and data collection. Produced all images, tables and manuscript.	
*Steven W. Polyak	Performed all recombinant DNA cloning including construction of expression vectors. Optimised the purification protocol for SaBPL and produced reagents required for purification of apo SaBPL. Performed enzymatic analysis and assisted with manuscript preparation.	
*Grant W. Booker	Provided intellectual discussion and assisted in manuscript preparation.	
*John C. Wallace	Provided intellectual discussion and assisted in manuscript preparation.	
Matthew C. Wilce	Provided direction and assistance in all aspects of X-ray crystallography, data collection, analyses and manuscript preparation and editing.	

* This work was supported through a Commercial Development Initiative grant from BioInnovationSA awarded to SWP, GWB and JCW.

CRYSTAL STRUCTURES OF APO AND LIGANDED BIOTIN PROTEIN LIGASE FROM *STAPHYLOCOCCUS AUREUS:* TOWARDS THE DEVELOPMENT OF NEW ANTIBIOTICS FOR MRSA

Nicole R. Pendini[1,2], Steven W. Polyak[1], Grant W. Booker[1], John C. Wallace[1], & Matthew C. J. Wilce[2*]

School of Molecular and Biomedical Science[1], University of Adelaide, North Tce, Adelaide, South Australia, Australia 5005, and, Department of Biochemistry and Molecular Biology[2], School of Biomedical Sciences, Monash University, Clayton, Victoria, Australia 3800.

Running head: Structure of Biotin Protein Ligase from Staphylococcus aureus

Address correspondence to: Matthew CJ Wilce, Building 13D, Department of Biochemistry and Molecular Biology, Monash University, Victoria, Australia, 3800, Tel: 613 9905 1086; Fax: 613 9905 3726; Email: matthew.wilce@med.monash.edu.au

Antibiotic resistance is becoming a serious problem facing human health. *Staphylococcus aureus* has developed resistance to most front line antibiotics available leading to many severe disease states. Biotin protein ligase (BPL) is responsible for the formation of the active form of biotin (biotinyl-5'-AMP) that activates essential carboxylases required in lipid biosynthetic pathways. To evaluate the potential of *Staphylococcus aureus* Biotin Protein Ligase (SaBPL) as a novel target for antibiotic drug design, the crystal structures of both apo and liganded SaBPL have been determined. The crystal structures indicate that SaBPL undergoes a monomer to dimer transition upon binding either biotin or biotinyl-5'AMP and unlike its *E. coli* counterpart, SaBPL does not undergo a 'disorder-to-order' transition when bound to biotin and then biotinyl-5'-AMP. Furthermore, the positioning of the dimer interface changes 37^{o} when compared to EcBPL. The catalytic domain positions the winged helix motif at the N-terminal domain for DNA binding and is consistent with bioinformatics studies of key DNA residues in the operon. The electrostatic surface potential for SaBPL and the biotin acceptor domain suggest a completely different mode of transfer of biotin to that proposed for *P. horikoshi* BPL. Comparison of SaBPL, the first BPL structure from a gram-positive microorganism, to known BPL structures identifies non-homologous residues around the ligand-binding site and we propose these differences make SaBPL an attractive target for novel antibiotic development.

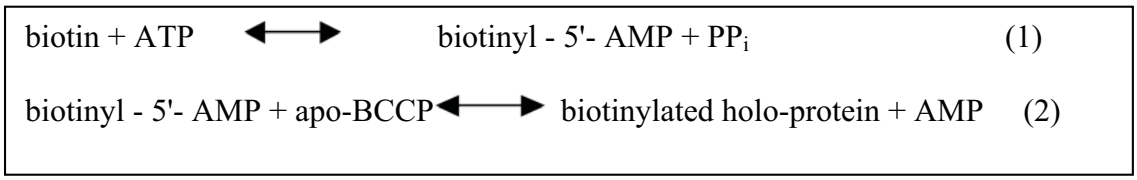
Introduction

Biotin protein ligase (BPL)

Biotin protein ligase (BPL) is the enzyme responsible for the posttranslational modification of target proteins with the essential prosthetic group, biotin. BPL ligates biotin onto specific carboxylases, transcarboxylases and decarboxylases within the cell and is crucial in both eukaryotes and prokaryotes [1]. These biotin-requiring enzymes are involved in essential metabolic pathways such as gluconeogenesis [2], lipogenesis [3], amino acid metabolism [4,5] and energy transduction[6] highlighting their fundamental importance in cellular processes [7]. BPL catalyses the nucleophilic attack on the α-phosphate group of ATP by the oxygen of the biotin carboxylate group, leading to the production of the active species biotinyl-5'-AMP (reaction 1). This reaction is analogous to the lipoate protein ligase (LplA) reaction and it is thought that both BPL and LplA have evolved from the class II tRNA synthetase protein lineage [8]. Once BPL has produced biotinyl-5'-AMP it can catalyse the covalent attachment of biotin onto a specific lysine residue that resides in the BCCP domain of the target protein.

The crystal structures from five different sources of BPLs have been determined. The 3-dimensional structure of apo BPL from *E. coli* (EcBPL) was the first to be determined by x-ray crystallography [9]. The structure of EcBPL in complex with biotin as well as with the non-hydrolysable biotinyl-5'-AMP analogue, biotinol-5'-AMP (BtOH-AMP) have also been determined [8,10]. EcBPL (also know as BirA - Biotin inducible repressor) is composed of three domains including an N-terminal DNA binding domain, a central catalytic domain and a C-terminal cap domain. EcBPL has been shown to be a bifunctional protein, with both enzymatic and autoregulatory functions [11]. EcBPL was thought to bind biotin before MgATP based on affinities for each ligand [12] but subsequent structural studies uncovered that biotin triggers ordering of the ATP binding site to allow binding [10]. Upon synthesis of biotinyl-5'-AMP (also known as the corepressor) EcBPL undergoes homodimerisation which facilitates the binding of BPL to the biotin biosynthetic operon thereby preventing *de novo* biotin biosynthesis [13]. *E. coli* only possess one biotin requiring enzyme, namely Acetyl CoA Carboxylase (reaction 2) [14].

The crystal structure of BPL from the archaeon *Pyrococcus horikoshi* OT3 (PhBPL) has been determined, revealing no DNA binding motif at the N-terminus and therefore does not share an analogous autoregulatory role to EcBPL. Structures of PhBPL have been determined in the apo form and with the ligands biotin and biotinyl-5'-AMP [16]. Recently, the crystal structure of PhBPL in complex with its substrate, the apo-biotin carboxyl carrier protein (BCCP) of methylmalonyl CoA decarboxylase, has been solved too [17]. The structure of PhBPL in complex with BCCP reveals important residues that are involved in the biotinylation reaction that are conserved in SaBPL including Trp101-Asp104, Leu116 and Glu118 (Trp177-Asp180, Leu192 and Glu194 SaBPL numbering). The crystal structures of other BPLs including *Methanococcus jannaschii* BPL (biotin bound) and from *Aquifex aeolicus* and *Mycobacterium tuberculosis* in unliganded form[18] have also been determined by x-ray crystallography and deposited in the protein data bank.



The importance of Staphylococcus aureus BPL

Staphylococcus aureus is a gram-positive bacterium that can be both a commensal and a pathogenic organism. This organism has many mechanisms that assist in its ability to evade the host immune system including production of toxins [19-21], ability to adhere to plasma proteins and survive within many host cell types [22]. This results in disease states such as septic and toxic shock, food poisoning, skin lesions, bloodstream infection and death.

Antibiotics have routinely been used to control *S. aureus* infection. The first line antibiotic, penicillin, was developed and used against *S. aureus* in the 1940's. Since this time, many alternative antibiotics have been discovered and found useful including methicillin, vancomycin, linezolid, tigecycline, and daptomycin [23-25]. *S. aureus* has evolved and developed resistance to all these once useful drugs, many after only two years of their wide spread use [26]. It has been shown that exposure to these antibiotics initiates an "SOS response" in bacteria, leading to upregulation of DNA repair proteins, thereby allowing horizontal transfer of resistance or virulence genes to offspring or through cell-to-cell contact between various bacterial species [27]. This has lead to the emergence of multi-drug resistant *S. aureus* isolates. New generation antibiotics in development are structural relatives or derivatives of those already on the market, exerting their effect on analogous pathways to the parent compounds. These are failing to stop multi-drug and methicillin resistant *Staphylococcus aureus* (MRSA) [28]. More recently vancomycin resistant *S. aureus* (VRSA) have been identified and combated with Linezoid, though in 2001 the first resistant *S. aureus* to this final line of defence was clinically identified [29].

Due to their fundamental importance for bacterial cell survival and maintenance, enzymes involved in the fatty acid synthesis pathway are of interest as potential anti-bacterial drug targets. One of the most successful anti-tuberculosis drugs, isoniazid, inhibits elongation of fatty acid by inhibition of NADH-dependent enoyl reductase [30]. Anti-bacterial compounds targeting ACC have also been identified and are being tested against *S. aureus* [31,32]. To date no drugs have been developed to target BPL though this concept underpins recent reports of the structure and biophysical characteristics of BPL from *M. tuberculosis* [33]. Indeed BPL has been shown to be an essential enzyme in bacteria, with mutations (particularly to the active site) or truncations to EcBPL inhibiting function and growth of the organism [34-36].

To aid in the development of inhibitors to SaBPL, and the discovery of new antibiotics against *S. aureus* we have determined the three dimensional structure of BPL from *Staphylococcus aureus* (SaBPL). Targeting SaBPL will lead to debilitating the function of both ACC and PC, impeding several pathways including lipogenesis and anaplerosis which are critical for cellular survival. Here we present the first structure of a gram-positive bacterial BPL in complex with biotinyl-5'-AMP (holo-SaBPL) as well as complexed with biotin (Bt-SaBPL), and in unliganded (Apo-SaBPL) form. These SaBPL structures reveal intrinsic differences between SaBPL and BPLs from *E. coli* and *P. horikoshii*.

Experimental Procedures

Production of SaBPL in complex with biotin (bt-SaBPL) or biotin-5'-AMP (*Holo-SaBPL).*
The cloning, overexpression and purification of recombinant SaBPL as well as the crystallisation has been previously reported [37].

Production of Apo SaBPL (Apo-SaBPL).
Whilst the displacement of biotinyl-5'-AMP from holo-SaBPL could be readily achieved by incubation with a large excess of biotin, its removal to obtain the apo-SaBPL was not so straightforward. To achieve this, cell lysate containing overexpressed SaBPL was initially incubated with a BPL substrate for 1 hour at 37^{o}C prior to purification. Here the apo-biotin form of *E. coli* biotin carboxyl carrier protein (EcBCCP-87) was employed to accept the biotin moiety. EcBCCP-87 was prepared as a fusion to GST to facilitate the production and purification of the substrate. Firstly an expression construct was prepared by PCR amplification of the coding sequence using oligonucleotides B1 (ATCTACGGATCCATGGAAGCGCCAGCAGCAGC) and B2 (ATCTACGAATTCATCACTCGATGACGACCAGCGG) together with plasmid pTM53 [38] as template. The primers were designed with 5' BamH1 and 3' EcoR1 restriction sites respectively, facilitating cloning into pGEX-4T-2 (GE Healthcare). The resulting construct, pGEX-BCCP87, was introduced into *E. coli* BL21 cells for protein expression. Expression and purification of the GST fusion protein was performed essentially as described [38]. After purification, the lysate mix was then used to purify apoSaBPL as described for holo-SaBPL. The presence of apo enzyme was detected by native PAGE on 12 % polyacrylamide gels to be > 95 % pure and was stored at 2 mg mL^{-1} in storage buffer (50 mM Tris-HCl pH 7.5, 0.5 mM EDTA pH 8.0 and 250 mM NaCl) at -80^{o}C. The protein was then concentrated to 13 mgmL^{-1} and crystals of apo-SaBPL were grown in 12 % PEG 8000 and 0.1 M Tris-HCL pH 7.5 at 277 K in a 1:1 ratio using the hanging-drop vapour diffusion method.

Structure Determination of SaBPL.
X-ray diffraction data for holo-SaBPL was collected in house on a Rigaku RUH2R X-ray source with a rotating copper anode equipped with Osmic confocal optics, an R-Axis IV detector, and an Oxford Cryosystems 700 Series cryostream. Diffraction data for Bt-SaBPL and apo-SaBPL were collected at the high-throughput protein crystallography beamline at the Australian Synchrotron using a MAR research 165 CCD detector. The diffraction data were integrated using MOSFLM [39] and the intensities were merged and scaled using SCALA [40]. Wilson scaling was applied using TRUNCATE [40]. Initial phases were determined by molecular replacement using PHASER with the apo *Pyrococcus horikoshii* OT3 BPL structure, PDBID: 1WQ7, as the search model in CCP4 [40]. Residues 71-324 of Holo-SaBPL were able to be interpreted in the electron density maps. PIRATE was employed to improve the

phases, followed by BUCANNEER, which built in approximately 90% of the missing N-terminal residues. The model was built using cycles of manual model building using COOT and refinement with REFMAC [41]. The partial occupancy of the adenosine phosphate moiety of biotinyl-5'-AMP was estimated by varying the occupancy until the refined B-factors approximated that of the rest of the ligand and the side chains adjacent to the adenine group. The quality of the final models were evaluated using MOLPROBITY [42]. Statistics for refinement and percentage of residues in the allowed region of the Ramachandran plot have been reported in Table 1. Figures 1a,b,c, 2, 3b, 4 and Supplementary Figure 1 and 2b were prepared using PYMOL [43], sequence based structure alignment (Figure 3a) using ESPript [44]. Figure 6 was generated using CHIMERA [45]. Ligplot was used to generate supplementary Figure 1d. Surface areas were calculated using PISA [46].

Modelling of the SaBPL dimer on the Biotin Biosynthetic Operon

Coordinates for B-form DNA with the sequence of biotin synthetic operon BPL binding motif were produced using the program B [47]. The positioning of DNA relative to the DNA binding domain of SaBPL was achieved by superposing one subunit of the winged helix protein RTP complexed to dsDNA (PDBID:1F4K [48]) using LSQMAN [49], Co-ordinates to B-form DNA with the operon sequence were produced using B [47] , and Chimera [45] was employed to superpose the B-form DNA of the biotin synthetic operon onto the DNA of RTP.

Results and discussion

Overview of SaBPL structure determination.

Recombinant SaBPL was overexpressed in *E. coli* BL21(DE3) cells and purified to a high level of homogeneity (> 95 %) using nickel affinity followed by cation exchange chromatography [37]. Three different forms of SaBPL were crystallised: SaBPL:biotinyl-5'-AMP (holo-SaBPL), SaBPL:biotin (bt-SaBPL) and unliganded-SaBPL (apo-SaBPL). The production of apo-SaBPL required the removal of biotin-5'-AMP which was achieved by incubating holo-SaBPL with the apo-biotin carboxyl carrier protein (BCCP) from *E. coli* to which SaBPL has been previously shown to be able to transfer its biotin group [37].

The crystal structures of holo-SaBPL and bt-SaBPL were refined to 2.6 Å and 3.2 Å resolution respectively. These ligand-containing SaBPL forms crystallise as homo-dimers and share very high structural similarity (a root mean square deviation: R.M.S.D of 0.376 Å for 322 C^{α} atoms). The electron density was clearly interpretable for almost all residues of holo-SaBPL (with the exception of Met1 and residues Asp221 to Arg225 for which density was evident for the main chain but not readily interpretable for the side chains of these residues). Electron density was also clearly interpretable for most residues of apo-SaBPL with the exception of the entirety of residues Met1 and Lys 118 to Ser 129. The dimeric state of the ligand-containing SaBPL forms is consistent with their behaviour in solution as characterised by native PAGE analysis (results not shown). In contrast, the crystal structure of the apo form of SaBPL (apo-SaBPL), which was refined to 2.3 Å resolution, is monomeric. Apo-SaBPL also shows the greatest structural differences from the liganded SaBPL forms.

Table 1: Crystallographic data and refinement statistics

	Apo-SaBPL	Bt-SaBPL	Holo-SaBPL
Wavelength (Å)	0.95364	0.99810	1.54
Resolution range (Å)	35 – 2.10 (2.1- 2.15)	19.80 - 3.23 (3.23- 3.31)	20- 2.60 (2.6- 2.67)
Total observations	164316	88174	351340
R-factor (%)	19.59	20.18	19.93
R-free (%)	24.9	26.78	25.69
* R-merge	11.3 (53.3)	14.3 (49.1)	5.7 (36.2)
DPI (based on free R factor)	22.60	47.42	27.58
Unique reflections	17599 (1295)	9332(634)	18509
Completeness (%)	96.06 (96.3)	98.68 (96.8)	98.25 (94.7)
Redundancy	3.8 (3.9)	9.6 (9.8)	4.1 (3.7)
Mean $I/\sigma(I)$	8.8 (2.8)	8.1 (2.0)	12.7 (2.1)
Waters	313	28	91
r.m.s.d			
Bond lengths (Å)	0.019	0.010	0.022
Bond angles (deg.)	1.95	1.31	1.96
Ramachandran plot			
Most favoured (%)	96.0	80.5	89
Additional allowed (%)	5.3	17.5	10
Generously allowed (%)	-	1.7	0.7
Disallowed (%)	0.7	0.3	0.3
Unit cell parameters	a=50.13 Å, b=51.4 Å, c=67.58 Å ($\alpha=\gamma=90^o$, $\beta=108^o$)	a= 94.27 Å, b=94.27 Å, c=130.89 Å ($\alpha=\beta=\gamma=90^o$)	a=93.56 Å, b=93.56 Å, c=130.65 Å ($\alpha=\beta=\gamma=90^o$)
Space group	$P\ 2_1$	$P\ 4_2\ 2_1\ 2$	$P\ 4_2\ 2_1\ 2$

* R-merge = $\sum_{hkl} \sum_j |I_j(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_j \langle I(hkl)\rangle$, where $I_j(hkl)$ and $\langle I(hkl)\rangle$ are the observed intensity of measurement *j* and the mean intensity of the reflection with indices *hkl*, respectively.

The SaBPL structures all conform to the same overall fold as previously been reported for EcBPL structures [9,16]. The monomeric unit can be divided into 3 discrete domains (Figure 1a). The N-terminal domain (residues 2-61) consists of a winged helix-turn-helix motif consistent with a DNA binding structure. The central catalytic domain (residues 68- 274) forms an α-β domain containing 8 β-sheets that form the backdrop for the active site and 7 α-helices that buttress against the β-sheet scaffold. Finally, the C-terminal domain (residues 282 - 323) is characterised by having a fold similar to an SH3 domain [50] and caps the active site. Two short random coil linkers (residues 62- 67 and 275–281) reside in between these three domains.

Figure 1

a

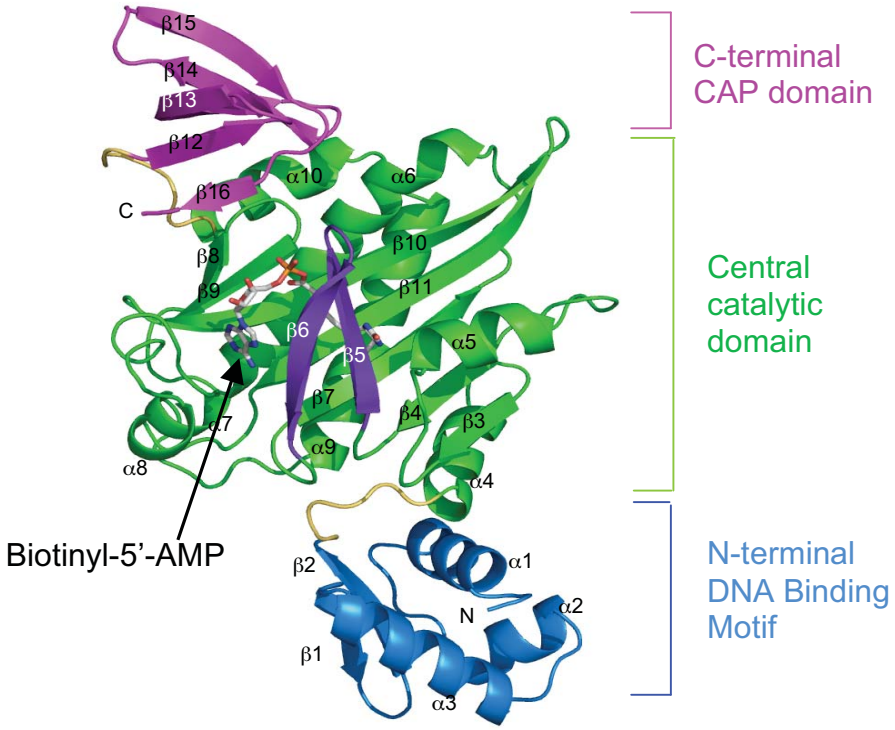


Figure 1a: Cartoon representation of Holo-SaBPL. (a) N-terminal domain residues 2-60 (blue), catalytic domain residues 68-274 (green) and C-terminal cap residues 282 -323 (magenta). Linker regions 62-67 and 275- 281 between these domains are shown in yellow. The reaction intermediate, biotinyl-5'-AMP, is shown by stick representation.

SaBPL crystallises as a stable dimeric holo-enzyme (holo-SaBPL)

Examination of the electron density revealed that biotinyl-5'-AMP was present in the active site and that liganded SaBPL exists as dimer (Figure 1b and 1c). The presence of biotinyl-5'-AMP was unexpected as attempts to determine the structure of *E. coli* BPL containing biotinyl-5'-AMP were not possible due to the rapid hydrolysis of the mixed anhydride upon X-ray exposure [8]. The data here shows that, in the case of SaBPL, the biotinyl-5'-AMP is maintained during over-expression and purification and that the biotinyl-5'-AMP is not severely affected by exposure to X-rays. However partial occupancy of the adenosine-phosphate component of the ligand (0.5) in the holo-SaBPL structure was observed, suggesting there may be partial degradation or disorder on exposure to X-rays.

b

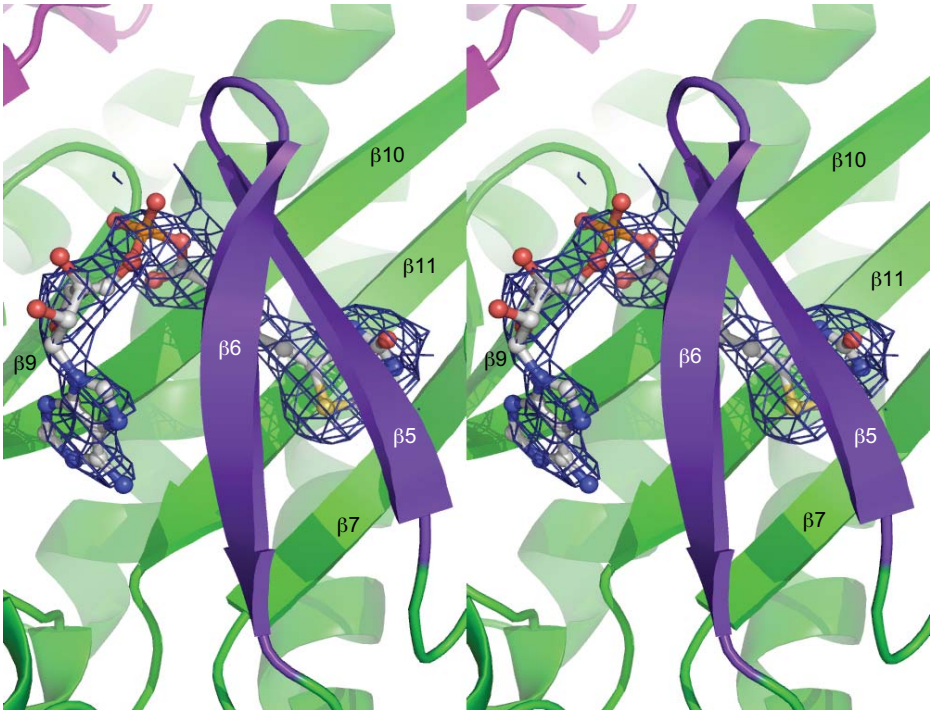


Figure 1b: A stereo image of the active site SaBPL is shown at 2.6 Å resolution (in stereo). The final $2F_o$-F_c map is contoured at the 1 σ level shown on biotinyl-5'-AMP (ball and stick representation).

c

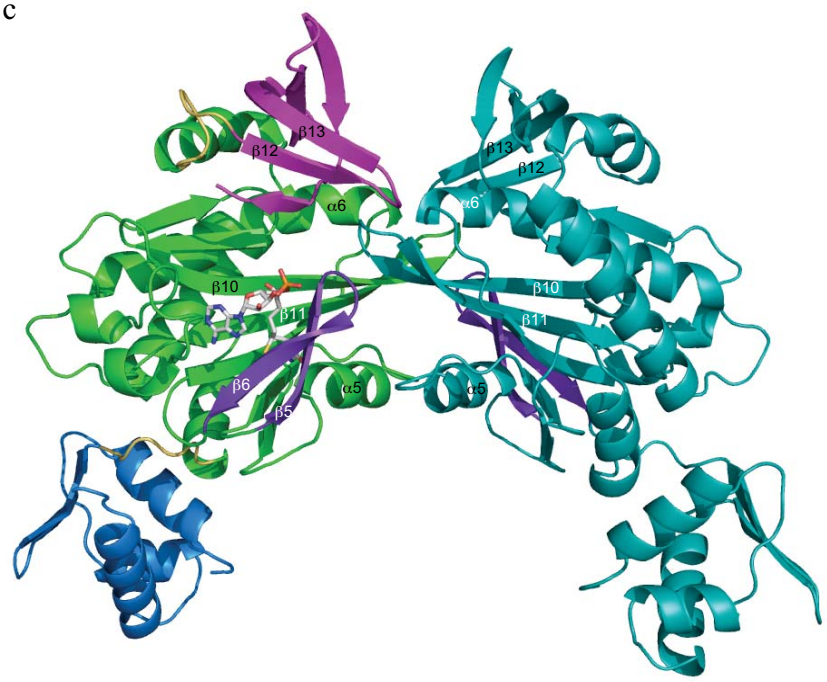


Figure 1c: shows the holo-dimer of SaBPL with structural elements involved in dimerisation labelled.

The holo-SaBPL structures shows that biotinyl-5'-AMP binds in a tight "U" shape geometry in a relatively open binding pocket. The biotin-binding site of SaBPL consists of a hydrophobic cavity or cave with a number of coulombic interactions with the polar ureido head group of biotin (Figure 1d) arising from residues outside the cave including Ser93 and Thr94. The amine functional groups of Arg122 and Lys187 interact with the carboxyl group of the valeric acid tail. Noticeably the ureido group is exposed to the bulk solvent, as too is the majority of the adenosine and ribose ring of the AMP. Residues of β10 (including Lys187 to Phe191) and the loop between α7 to α8 (including Ile224 and Ala228) form the AMP binding pocket while β5, β6 and the loop between these (residues Arg122 to Asn124) act like a similarly shaped "U", pressing over the reaction intermediate. Trp127 positions the adenylate moiety via π-π stacking interactions. The conformation of the ligand (corepressor) when bound to SaBPL positions the carbonyl carbon atom of biotin ready for nucleophilic attack by the target lysine residue of BCCP and may help the reaction by inducing strain in the bonds about this component of the ligand.

The dimerisation interface contains an extensive network of hydrogen and ionic bonds (summarised in Supplementary Tables 1 and 2). The C-terminal and catalytic domains orientate themselves to make contact with the analogous residues on the other subunit including interplay between α5, α6, β10 and the loops between β10-β11 with the active site loop, β12-β13 loop and α6 (Figure 1c). The salt bridges formed by Arg122 to Asp200 on the opposite subunit immobilise the active site loop, containing Arg122, preventing enzymatic function and release of the reactive biotin intermediate. The N-terminal domains do not make contact and are spaced approximately 36 Å apart at their closest point (Lys3).

Supplementary table 1: Hydrogen bonds between the dimer interface of Holo-SaBPL.

##	Chain A	Dist. [Å]	Chain B
1	A:LYS 100[NZ]	2.7	A:GLU 203[OE2]
2	A:ARG 122[NH1]	3.6	A:ASP 200[OD1]
3	A:ARG 122[NH1]	3.4	A:ASP 200[OD2]
4	A:SER 151[OG]	3.4	A:ASN 199[OD1]
5	A:MET 195[N]	2.8	A:ALA 197[O]
6	A:ALA 197[N]	3.0	A:MET 195[O]
7	A:ASN 198[ND2]	3.4	A:GLU 194[OE2]
8	A:SER 318[OG]	3.4	A:ASP 200[OD2]
9	A:GLU 203[OE2]	2.7	A:LYS 99[NZ]
10	A:ASP 200[OD1]	3.6	A:ARG 122[NH1]
11	A:ASP 200[OD2]	3.4	A:ARG 122[NH1]
12	A:ASN 199[OD1]	3.4	A:SER 151[OG]
13	A:ALA 197[O]	2.8	A:MET 195[N]
14	A:MET 195[O]	3.0	A:ALA 197[N]
15	A:GLU 194[OE2]	3.4	A:ASN 198[ND2]
16	A:ASP 200[OD2]	3.4	A:SER 318[OG]

Supplementary table 2: Salt Bridges between the dimer interface of Holo-SaBPL

##	Chain A	Dist. [Å]	Chain B
1	A:LYS 99[NZ]	3.4	A:GLU 203[OE1]
2	A:LYS 99[NZ]	2.7	A:GLU 203[OE2]
3	A:ARG 122[NH1]	3.6	A:ASP 200[OD1]
4	A:ARG 122[NH1]	3.4	A:ASP 200[OD2]
5	A:GLU 203[OE1]	3.4	A:LYS 99[NZ]
6	A:GLU 203[OE2]	2.7	A:LYS 99[NZ]
7	A:ASP 200[OD1]	3.6	A:ARG 122[NH1]
8	A:ASP 200[OD2]	3.7	A:ARG 122[NH1]

d

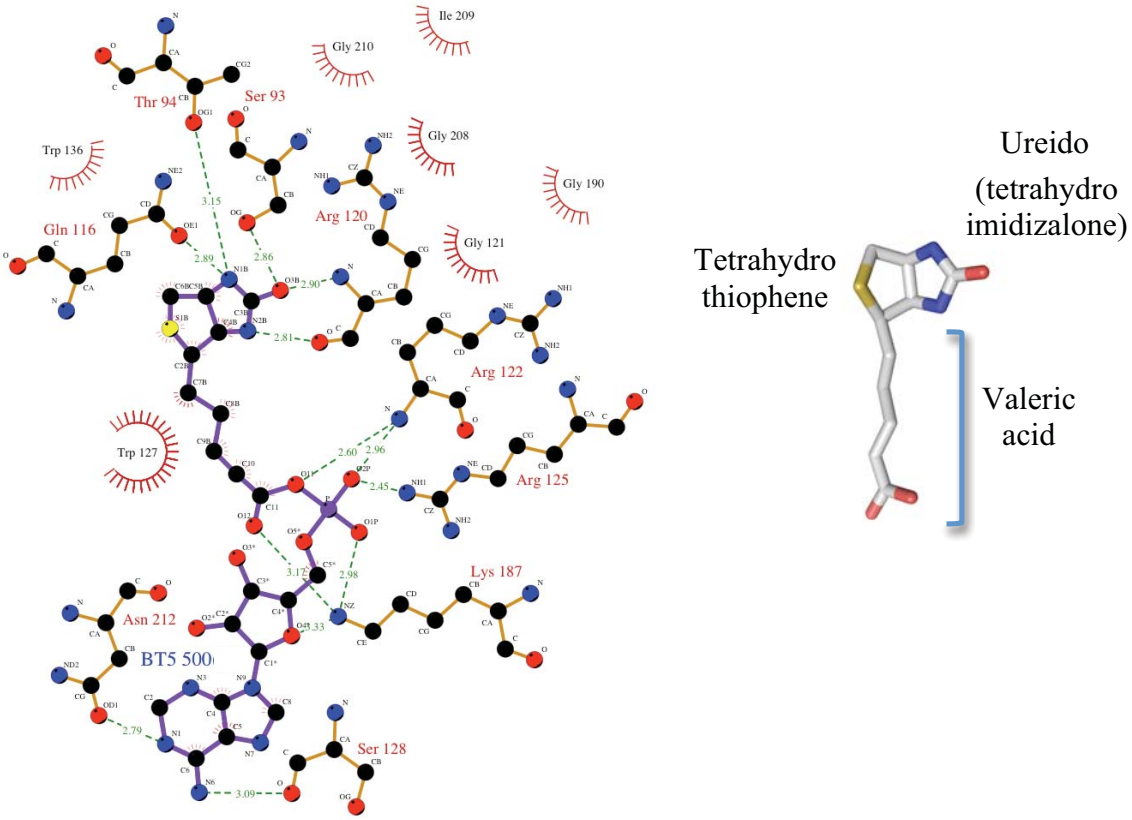


Figure 1d: A ligand plot of the bonding interactions between SaBPL and biotinyl-5'-AMP. Residues involved in hydrogen bonding are shown in orange and hydrophobic interaction in red. Biotinyl-5'-AMP is shown in purple and hydrogen bonding distance in Å by green dashed lines. To the right hand side is the structure of biotin with functional groups labelled.

Structural details of SaBPL in complex with biotin (bt-SaBPL)

The bt-SaBPL complex was produced by addition of 10-molar excess biotin to the protein prior to crystallisation. Electron density confirmed the presence of only biotin, there is no electron density visible, even at low contour levels, for adenosine or phosphate. Residues involved in hydrophobic interactions between SaBPL and biotin are analogous to those seen for the biotinyl-5'-AMP as too are the hydrogen bonding interactions (summarised in Table 2). The structure shows that the enzyme forms a dimer even when biotin alone is present and the dimer is isomorphous to the holo-SaBPL dimer. However, some of the interactions between the dimer interface have changed when compared to biotinyl-5'-AMP bound SaBPL dimer, including 2 additional salt bridges formations that include Arg122 [NH2] to Asp200 [OD2] (3.65 Å) and Asp200 [OD2] to Arg122 [NH2] (3.23 Å) and the loss of 3 hydrogen bond interactions, namely Arg122 [NH1] to Asp200 [OE2] (3.64 Å), Ser151 [OG] to Asn199 [OD1] (3.35 Å) and Asn199 [OD1] to Ser151 [OG]. In addition, the rotational variation of the dimer interface that was observed for the EcBPL biotin bound versus biotinol-5'-AMP bound is not seen for SaBPL.

Apo-SaBPL crystallises as a monomer with a disordered loop.

The formation of apo-SaBPL required the recombinant enzyme to be incubated with apo-Biotin Carboxyl Carrier Protein domain of *E. coli* ACC to discharge the biotinyl-5'-AMP from the active site. It has been previously shown that there is cross reactivity between BPLs and substrates from other species [51] and that EcBCCP can be used as a substrate for SaBPL in *in vitro* activity assays [37]. Apo-SaBPL was subsequently purified in the same manner as holo-SaBPL. The apo-SaBPL crystallised from conditions very similar to those of the liganded SaBPL. However the space group is different, namely $P2_1$ and the apo enzyme is observed to be a monomer. The structure has two disordered loops encompasing Thr116-Lys130 and Asp220-Arg224 where no electron density could be observed. The loss of structure in these loops effectively removes the "U" shaped press from the structure and completely opens up the ligand binding site ready to accept the biotin and ATP moieties.

Table 2: Comparison of hydrogen bonds between BPL and biotinyl-5'-AMP in *S. aureus, E. coli and P. horikoshii*.

SaBPL (Å)	EcBPL (Å) (NB: BtOH-AMP)	PhBPL (Å)	biotinyl-5'-AMP
Ser93 OG (2.86) Arg120 N (2.90)	Ser89 OG (2.8)	His46 N (2.99)	O3B
Thr94 OG (3.15) Gln116 O E1 (2.89)	Thr90 OG (3.1)	Thr22 OG (3.07) Gln 42 OE1 (2.84)	N1B
Arg120 O (2.81)	Arg116 O (2.7)	His46 O (2.97)	N2B
Lys187 NZ(3.12)	*	Lys111 NZ (2.7)	O12
Arg122 N (2.60)		Arg48 N (3.22)	O11
Arg122 N (2.96) Arg125 NH1 (2.45)	Arg118 NH1 (3.1)	Arg48 N (2.93) Arg48 NH1 (3.49)	O2P
Lys187 NZ (2.98)	Arg118 NH1 (2.7) Arg121 NH1 (3.1)	Arg48 NH1 (2.74)	O1P
Asn212 OD1 (2.79)	Asn208 NHD1 (2.7)		N1
Ser128 O (3.09)	Asn208 OD1 (3.1)	Glu54 O (2.98) Asn131 O (3.06)	N6

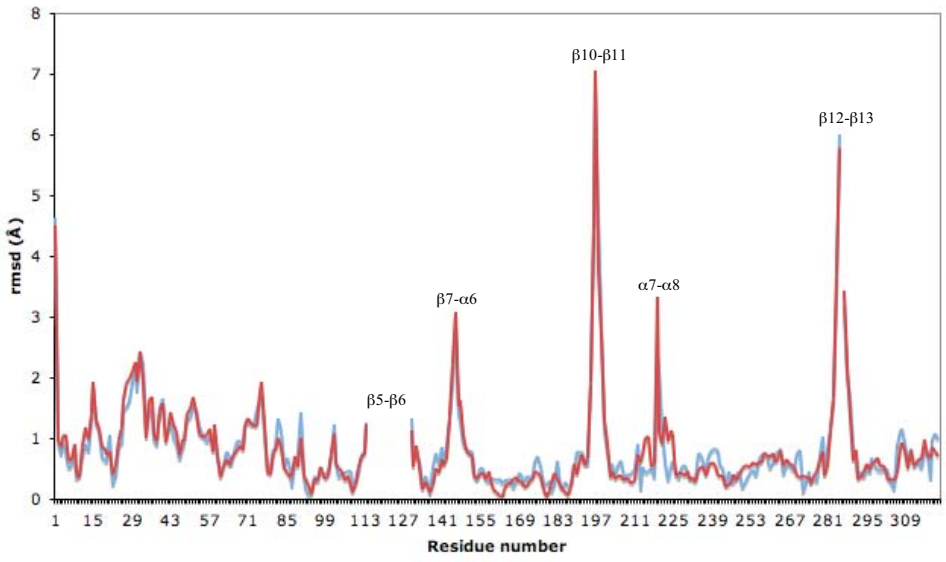
* NB this contact is not observed as the carbonyl group in biotinyl-5'-AMP is not present in BtOH-AMP.

Comparison between apo and ligand bound SaBPL

These structures reveal that there are conformational changes that take place when biotinyl-5'-AMP or biotin are bound at the active site compared to apo enzyme, particularly when the position of the main chain of each amino acid is compared. Relatively large local shifts are observed between the apo and ligand-bound SaBPL's including residues 116-132, 146-150, 197-202, 220-223 and 284-291 (These correspond to loops between β5-β6, β7-α6, β10-β11, α7-α8 and β12-β13 respectively, Supplementary Figure 1a). Residues 116-132 comprise the biotin binding loop which is not observed in the apo-SaBPL structure, and the mobility of this loop allows the binding of the substrates. Residues 146-150 pack against residues 197-202 allowing formation of a seamless β-sheet at the dimerisation interface with β7, β10 and β11. Residues 220-223 correspond to the solvent exposed loop over the adenosine-binding pocket. This loop positions Ile234 to form hydrophobic interactions with the adenine ring of AMP in the holo conformation. The loop is structured in bt-SaBPL whereas it is not in the holo structure, even though there is a disorder-to-order transition seen for the crystal structures of Bt-EcBPL to btOH-AMP-EcBPL. We propose the mobility of this loop is required to facilitate biotin-binding and, following the first partial reaction, shield the highly reactive mixed-anhydride bond in biotinyl-5'-AMP from the aqueous environment. The electron density for this loop suggests that more than 1 conformation exists, though at the resolution of these structures we are unable to confirm these with absolute certainty.

Supplementary Figure 1a

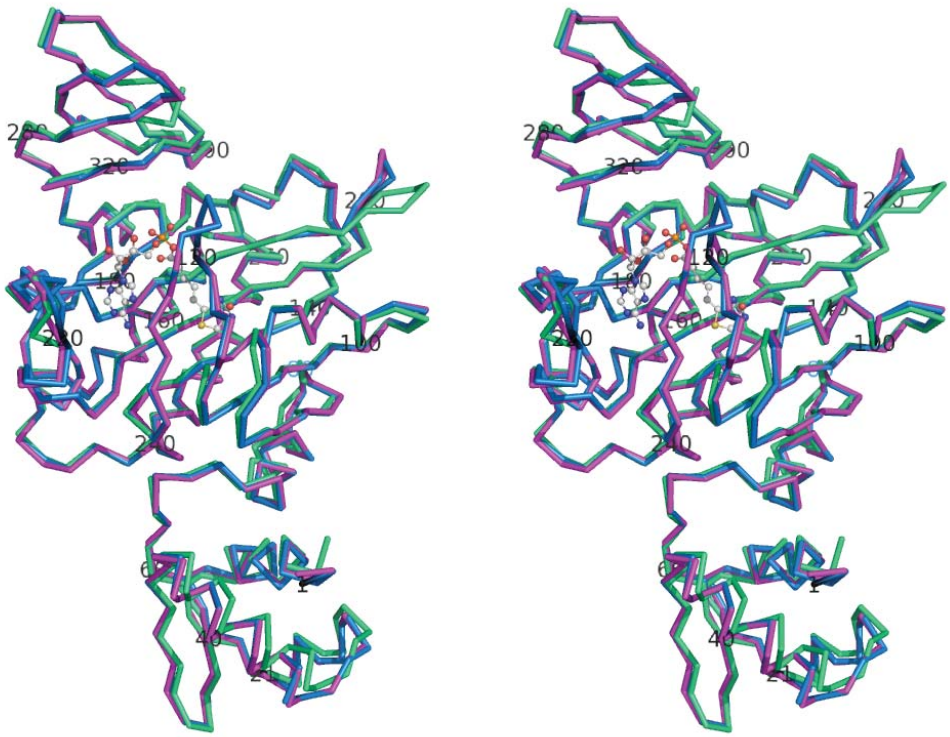
a



Supplementary Figure 1a: A plot of the main-chain movement between the unliganded forms and liganded forms of SaBPL after superposition. The R.M.S.D between main-chain residues are represented graphically in Angstroms where the biotin bound compared to apo SaBPL is drawn in blue and the biotinyl-5'-AMP bound SaBPL vs apo in red. Largest differences are between those residues in loop regions that are involved in ligand binding.

The global structural differences between apo and ligand bound SaBPLs are however minimal. This is highlighted when the C^{α} of apo-SaBPL are superposed on those of bt-SaBPL and holo-SaBPL which result in only a 0.876 (between 298 C^{α} atoms) and 1.013 Å (between 302 C^{α} atoms) R.M.S.D respectively (Supplementary Figure 1b). The C-terminal cap was initially undefined, but in comparing the liganded structures to the apo, a critical aspartic acid (Asp320) can be seen to form a salt bridge with an arginine (Arg125) in the catalytic core. This interaction stabilises the loops between β5-β6 and β14-β16 over biotinyl-5'-AMP, preventing its release from the enzyme. From the liganded structures, it would appear that this interaction is necessary to form the ligand binding site as a majority of the residues that conform the site are provided by the β-strands either side of this loop, namely β5 and β6. The importance of this interaction is seen with a point mutation to the equivalent residue, Arg317Glu in EcBPL, that causes a significant decrease in affinity for ATP [35].

b



Supplementary Figure 1b: Structural comparison between apo, biotin bound and holo SaBPL. Superimposed C^{α} traces represented by ribbons of unliganded (aqua), biotin bound (magenta) and biotinyl-5'-AMP bound (blue) SaBPL. The position of biotinyl-5'-AMP is shown in ball and stick representation, and every 20th amino acid of holo-SaBPL is printed.

Comparison of SaBPL to published BPL structures

Dimerisation comparison

When comparing the dimers of SaBPL and EcBPL (holo forms) it becomes obvious that there is a dramatic change in the orientation of the DNA binding domain, measuring 37.7^{o} around the crystallographic axis (domain axis) and a translation of 3.9 Å (DynDom) [52] (Figure 2a and 2b). The flexible linker between the N-terminal and catalytic domain consists of 7 residues for SaBPL, in comparison to 20 for EcBPL. EcBPL makes only minimal contact between the N-terminal and catalytic core, with hydrogen bonds between the linker region and the catalytic domain include Glu62 (OE2) to Arg212 (NH_2) (2.96 Å) and Leu66 (O) to Asn236 (NZ). In SaBPL these interactions can be observed between Gln61 (OE1) and Lys237 (N_2) (3.21 Å) and Gln5 (OE1) to Gly69 (N) in addition to the N-terminal and catalytic domain interaction Gln9 (NE2) to Gln73 (OE1) (2.76 Å), and the positioning of α1 much closer to α4, allowing more van der Waals interactions between the two domains. These internal domain interactions are important as the N-terminal domain is known to bind to the biotin biosynthetic operon in *E. coli* leading to repression of biotin biosynthesis. It is thought to have the same function in *S. aureus*, even though there is a major difference in the way the dimer forms leading to a change in the positioning of the N-terminal relative to one another.

Figure 2

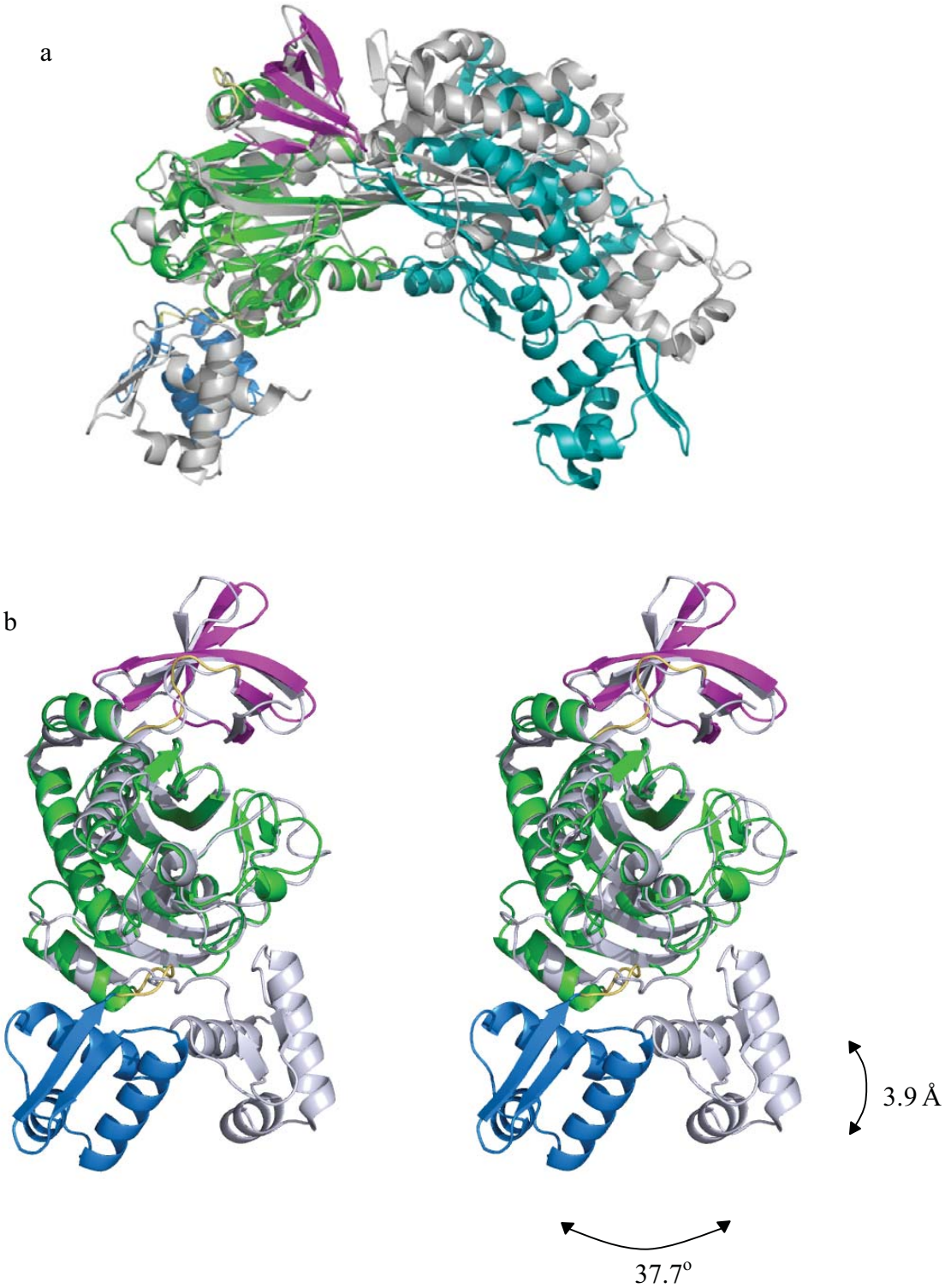


Figure 2: Dimerisation comparison between SaBPL and EcBPL.
(a) Ribbon representation of the dimer of SaBPL (coloured as per figure 1 and cyan) used to superposition on 224 residues of chain A- EcBPL (grey). (b) 90° rotation cartoon representation of (a) showing only chain A of SaBPL and EcBPL. Curved arrows in black show large differences between the 2 structures in the positioning of the N-terminal domain.

Catalytic domain

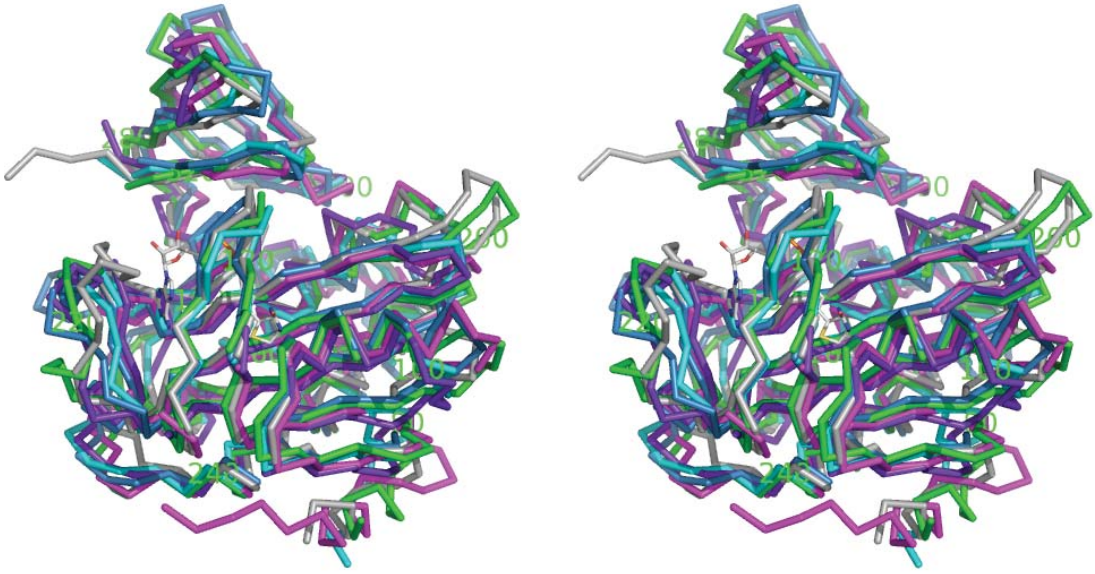
The structural similarity is very high between BPLs. This is particularly obvious when comparing the catalytic domain across all species that have had their structure determined. Superposition based on C^{α} positions for the catalytic domains of SaBPL, EcBPL and PhBPL, whether ligand bound or not, result in an overall R.M.S.D of approximately 2 Å between them (values for superimpositions between apo, biotin bound and holo BPLs are summarised in supplementary Table 3 and Supplementary Figure 2). For completeness, superposition has also been performed with the BPLs deposited in the PDB from *Methanococcus jannaschii* (biotin bound), *Mycobacterium tuberculosis* and *Aquifex aeolicus* (in apo form). Little more than the structure is known about the BPLs from these 3 species, and therefore the focus will be on comparing SaBPL to the *E. coli* and *P. horikoshi* BPLs which are the best characterized and have been published.

Supplementary table 3 :C^{α} superposition between different forms of reported BPLs (reported in Å).

	Apo-SaBPL	Bt-SaBPL	Holo-SaBPL
Apo-SaBPL			
Bt-SaBPL	0.9 (298)		
Holo-SaBPL	1.0 (302)	0.4 (322)	0.4 (322) AB
Apo-EcBPL (PDBID: 1BIA)	2.2 (217)	2.4 (225)	2.4 (226)
Bt-EcBPL (PDBID: 1HXD)	2.0 (218)	2.2 (236)	2.2 (236)
Holo-EcBPL (PDBID: 2EWN)	1.9 (226)	1.9 (245)	1.9 (249)
Apo-PhBPL (PDBID:1wq7)	1.8 (218)	2.0 (222)	1.9 (222)
Bt-PhBPL (PDBID:1wqy)	1.8 (211)	2.0 (226)	1.9 (226)
Holo-PhBPL (PDBID:1wqw)	1.8 (212)	1.9 (225)	1.9 (227)
Bt-*Methanococcus jannaschii* BPL (PDBID: 2EJ9)	1.9 (203)	2.0 (220)	1.9 (220)
Aquifex Aeolicus (PDBID: 2EAY)	1.9 (218)	1.9 (213)	1.8 (211)
Mycobacterium tuberculosis (PDBID: 2CGH)	1.9 (222)	2.1 (230)	2.0 (227)

NB: AB refers to subunit 1 and 2 of the dimer having been superpositioned.

Supplementary Figure 2



Supplementary Figure 2: Superimposed C^{α} traces of SaBPL (green), EcBPL (grey) and PhBPL (aqua), *M. Jannaschii* (blue), *M. tuberculosis* (magenta) and *A. aeolicus* (purple) catalytic and C-terminal domains. The position of biotinyl-5'-AMP in SaBPL is represented by stick model.

Ligand binding

Many of the residues important for ligand binding in SaBPL are highly conserved when compared to BtOH-AMP-EcBPL and Biotinyl-5'-AMP-PhBPL (as summarised in Table 2 and shown in a structure-based sequence alignment in Figure 3a). One notable difference between the binding of Biotinyl-5'-AMP is hydrogen bonding to O11 which is seen for both SaBPL (Arg122) (Figure 1d) and PhBPL (Arg48) but not in EcBPL, though the residue is conserved. This additional H-bond may account for the increased stability of the ligand in SaBPL and PhBPL allowing the holo structure of these BPLs to be determined by X-ray crystallography but not for EcBPL. The stability of biotin bound SaBPL may be due to resistance in X-ray damage and use of cryo-cooling or due to an additional H-bond with the ligand through Arg120, when compared to Bt-EcBPL. Residues that are involved in H-bonding and securing biotin in SaBPL are equivalent to those between EcBPL and biotin and for PhBPL and biotin.

Figure 3a

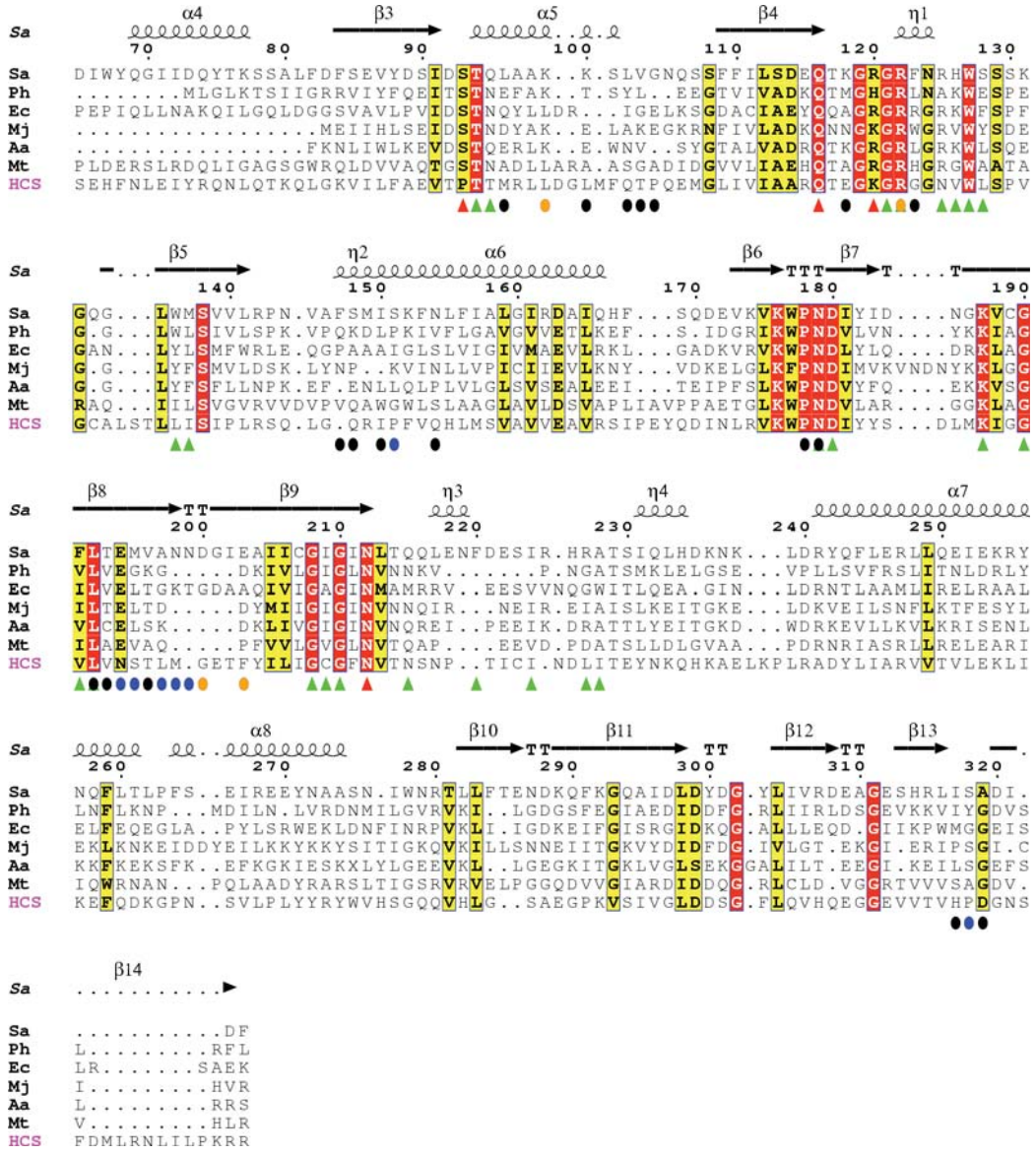


Figure 3a: Sequence alignment between structurally determined BPLs to human holocarboxylase synthetase. Red triangles indicate hydrogen bonding to Biotinyl-5'-AMP, green triangles indicate contact to Biotinyl-5'-AMP, orange circles indicate residues involved in salt bridge formation at the dimer interface, black circles indicate dimer hydrogen bonds and blue circles show dimer contact residues for Holo-SaBPL.

Comparison of the ligand bound in the active site of both SaBPL and EcBPL shows the ligand is more exposed in SaBPL. Furthermore, there is a small solvent exposed pocket over the oxygen of the uredio group in biotin for SaBPL, this is also seen for PhBPL (however smaller), yet for EcBPL this moiety is completely buried (Figure 3b). This is due to the repositioning of the side chain Arg120 in SaBPL (equivalent to Arg116 in EcBPL) that is in well-defined density in both SaBPL and EcBPL. In SaBPL the valeric acid end of the biotin-binding site is highly solvent exposed (the same is observed for PhBPL), however for EcBPL, Arg121 and Trp223 completely cover the site where AMP binds. These structures aid in explaining the differences seen in affinity for the ligands where the K_M for biotin in SaBPL is 1.1 ± 0.1 μM and 160 ± 19 μM MgATP. Whereas values reported for *E. coli* BPL using the same assay conditions resulted in a *Km* for biotin of 0.3 ± 0.1 μM and 200 ± 3 μM for MgATP [35]. This implies EcBPL has a slightly higher affinity for biotin than SaBPL. However close inspection of the structure of the binding site, SaBPL has no residues hindering the AMP binding site and therefore would favour the ligand-binding of ATP, observed by a lower K_M, than seen for EcBPL .

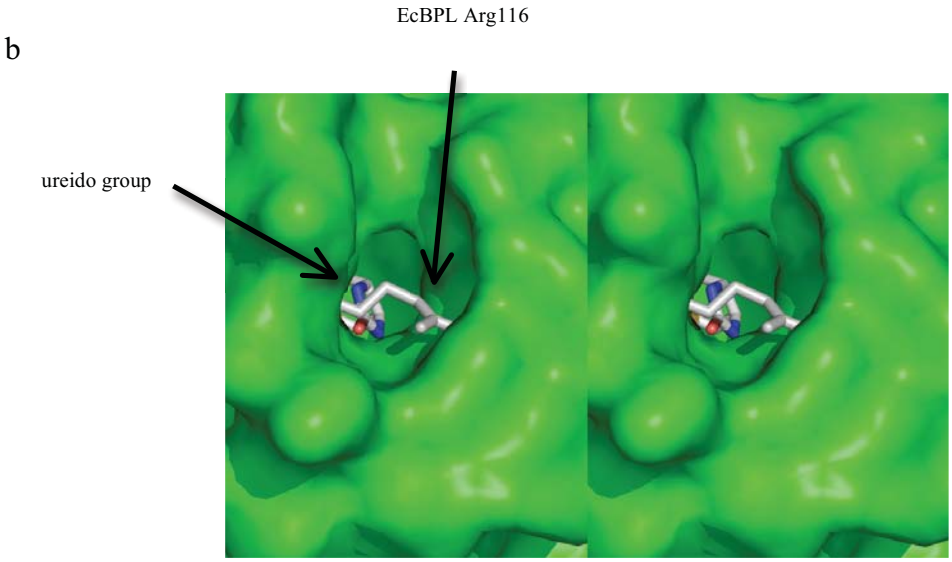


Figure 3b: Surface representation of the SaBPL active site showing the exposed ureido moiety of biotin. In EcBPL Arg116 covers the ureido group, burying the ligand completely.

Through superposition of the holo and biotin bound SaBPL we observe no significant movement of the N-terminal or the dimer interface between these forms, nor these to apo-SaBPL. Essentially, the unliganded to liganded forms of SaBPL are isomorphous. However large structural changes are observed between Bt-EcBPL and BtOH-AMP-EcBPL, where a shift of 19.2^{o} is observed between the catalytic domain to the N-terminal domain and a 12^{o} bend at the dimer interface [8] furthering the differences seen between the two BPLs. PhBPL forms a dimer interface with β3 (SaBPL numbering) whereby the structures sit on top of one another rather than a side-by-side interaction as seen for both EcBPL and SaBPL. *P. horikoshi* does not synthesise biotin *de novo* and that explains the lack of the DNA binding domain, but as to why this enzyme exists as a homodimer when it does not require repressor function remains unclear. One possibility is homodimerisation provides the stabilisation of the catalytic core that could be provided by the DNA binding domain in the bacterial BPLs.

The SaBPL structure differs from the previously crystallised BPLs as it co-purifies with the natural ligand, indicating it either has a higher affinity for the ligand or the dimer is more stable than that of EcBPL or both. The biotin bound EcBPL dimer interface covers 1168.8 Å of solvent-accessible surface area with 18 hydrogen bonds and 2 salt bridges, compared to 1020 Å of Bt-SaBPL which makes 13 hydrogen bonds and 10 salt bridges at the dimer interface. However due to the limited resolution of the Bt-SaBPL, it is difficult to draw many conclusions on this data and the positioning of the contact side chains. In the holo EcBPL dimer 1244 Å of solvent-accessible surface area is buried at the interface and 1029 Å for Holo-SaBPL. BtOH-AMP-EcBPL exchanges 16 hydrogen bonds between the interface and 4 salt bridges. Holo-SaBPL makes the same number of hydrogen bonds and has an additional 4 salt bridges at the dimer interface.

The structural information for the SaBPL dimer differs from that observed with the Bt-EcBPL to Bt-AMP-EcBPL transition where the holo dimer is a tighter dimer than the biotin-bound dimer based on buried surface area and decreased thermal factors. For ScBPL this transition was not observed. When the dimers of Biotinyl-5'-AMP SaBPL and Bt-SaBPL are superimposed, there is an R.M.S.D of 4.278 Å between 224 C^{α} of the catalytic domains of Chain A. The most evident change between the SaBPL and EcBPL is the movement of helices α5, which are positioned much closer together in SaBPL so Lys98 to Glu220 form a salt bridge between the dimer interface. This reinforces the network of salt bridges that are formed which are analogous to EcBPL (summarised in supplementary Tables 4 and 5) and provides a rationale why the SaBPL purifies as a ligand bound dimer. Alternatively this could be an artefact of recombinant production of SaBPL in *E. coli.* Consequently SaBPL may have converted all the apo substrate into holo, therefore with no further substrate to biotinylate, homodimerises.

Molecular mechanism of biotinylation

Recently the structure of PhBPL complexed with biotin carboxyl carrier protein (BCCP) subunit of methylmalonyl-CoA decarboxylase has been solved [17]. Superposition of this structure (PDBID 2ejf or 2ejg) to the homodimer of SaBPL shows that BCCP binds to PhBPL, positioning the exposed target lysine loop equivalent to the loop between β10 and β11 (residues 196-201) in the SaBPL dimer interface. Therefore, using the complex structure as a guide for the biotinylation mechanism in SaBPL, this would suggest that in order for enzymatic function to occur the affinity of Biotinyl-5'-AMP-SaBPL for the BCCP must be higher than it is for itself and thus no longer remains a homodimer. Therefore, competition or selectivity must exist between the biotin requiring enzymes and homodimerisation at the SaBPL "dimerisation interface", which has been suggested in the case of heterodimerisation for EcBPL [53].

Electrostatic and shape complementarity of molecular surfaces can be used as a method to predict how molecules interact or dock together. The electrostatic characteristics of the apo and holo structures of SaBPL and *S. aureus* Biotin domain (from Pyruvate carboxylase PDBID: 3bg5) have been calculated and the surfaces coloured according to electrostatic potential which were calculated with Coulombs' law in chimera (Figure 4). The holo-enzyme shows a positive region at the active site as well as more increased positive nature of the N-terminal for optimised interaction with negatively charged DNA. Of note the electrostatic surface of apo-SaBPL reveals no positive surface about the active site, and therefore would be unable to bind the BCCP substrate. From these observations it would appear that the negatively charged SaBCCP, approaches SaBPL on the opposite side of the dimerisation interface. This is in contrast to what is seen for the PhBPL- BCCP crystal structure suggesting there may be divergent mechanisms for biotinylation between species. We already see variation in the homo-dimerisation interface between archaeal and bacterial BPLs, and the same may be true of the heterodimerisation interface for BPL-BCCP.

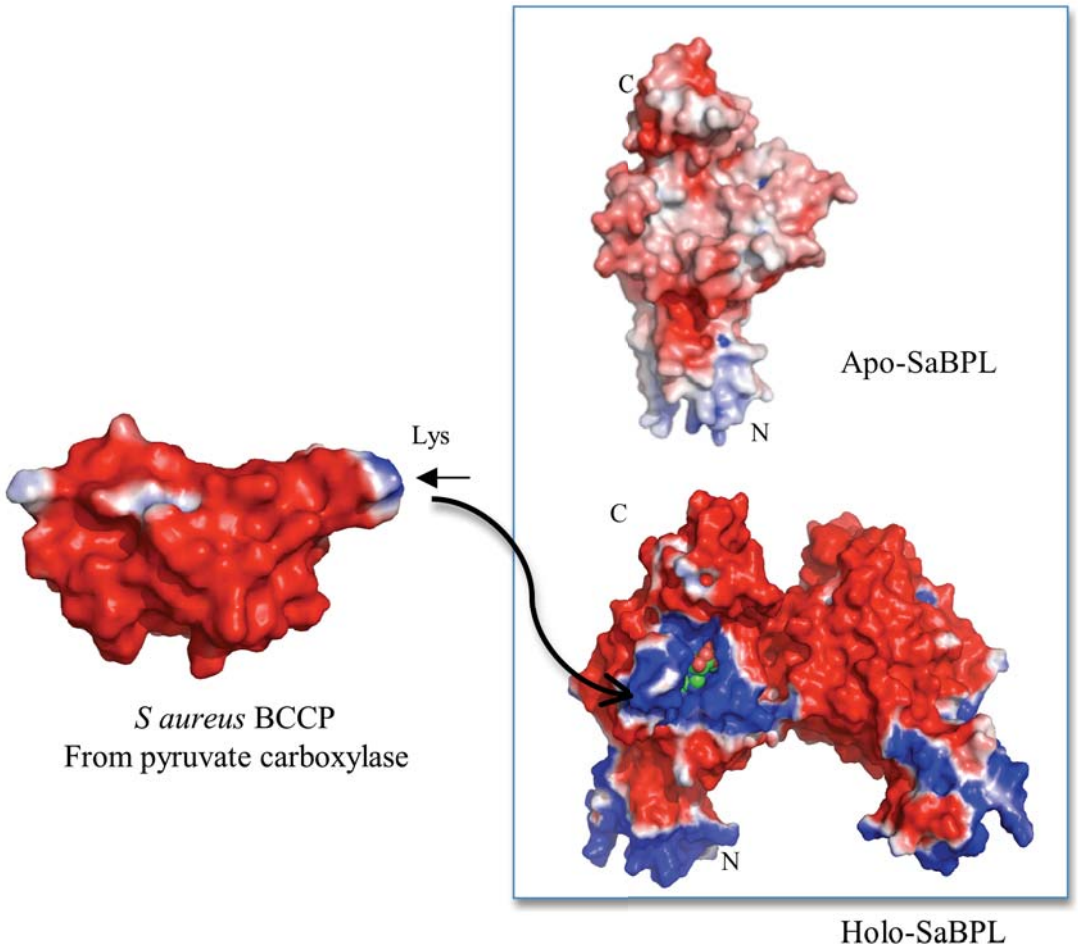


Figure 4: Electrostatics of SaBPL and *S. aureus* Biotin domain BCCP

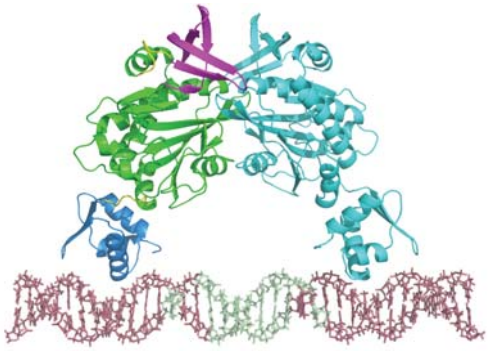
The structure of apo and holo-SaBPL shown in surface representation coloured red: indicating negative, white for neutral and blue for positive electrostatic characteristics. To the left is the surface representation of *S. aureus* EcBCCPiotin domain (from Pyruvate carboxylase PDBID: 3bg5PDBID: 2bdo), that is primarily negatively charged. Based on charge and shape, an arrow indicates that it potentially approaches from the opposite side to the dimer interface. The lysine that is biotinylated is indicated with an arrow, suggesting approach of the biotinylation domain to the opposite side of the dimer interface.

DNA binding and repression

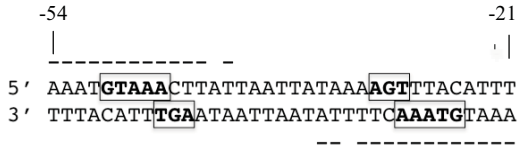
Binding of EcBPL to the biotin biosynthetic operon has been proposed and shown by DNaseI foot printing [54]. Rodionov *et al.* [15] have predicted residues that are involved in interactions with DNA, and these have been used to position SaBPL on to BioO. Modelling holo-SaBPL onto the biotin biosynthetic operon shows that helix 3 positions on the BioO sequence which lies 5' to the biotin biosynthetic genes bioD-bioA-bioB-bioF-bioW. The distance between helix 3 in the SaBPL dimer is approximately 66 Angstroms which approximates to the distance between the major grooves in B-form DNA (Figure 5). The distance between the EcBPL dimer is approximately 73 Angstroms, suggesting that either the dimer or DNA would require movement or bending to accommodate one another. The amino acid composition in helix 3 across BPLs shows high sequence conservation. This becomes important as the operon of BioD is shown to contain a pseudopalendromic sequence and helix 3 interacts with the same bases in the inverted repeat. Though these approximations fit known biological parameters, the difference between SaBPL and EcBPL juxta positions of the N-terminal may not be adequately determined until its structure is stabilised further and solved bound to DNA.

Figure 5

a)



b)



c)

Modelling the SaBPL dimer on the Biotin Biosynthetic Operon

Figure 5: a) Model of Holo SaBPL with *S. aureus* Biotin Biosynthetic Operon. The inverted repeat is coloured raspberry.

b) The sequence of the SaBPL binding region, positioned 21 nucleotides from the 5' end of bioD. (---) indicates the position of the inverted repeat. Bolded and boxed bases are in a position to interact with SaBPL.

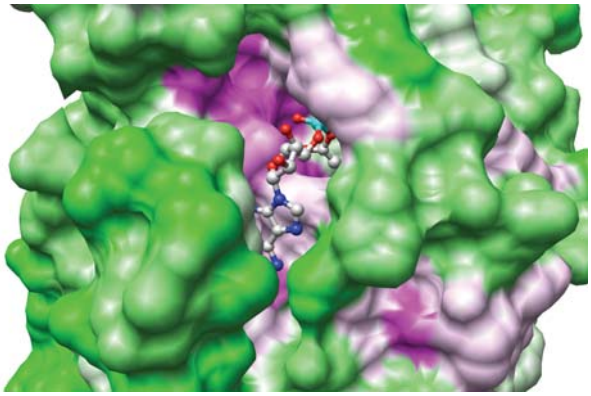
c) The sequence of part of the helix 3 that could interact with the DNA. The boxes indicate conserved residues in BPL winged helixes.

SaBPL as a drug target

Structures of BPL from *Staphylococcus aureus* have been solved unliganded, with biotin bound, and co-purified and crystallised with biotinyl-5'-AMP. We observe many differences between SaBPL and those structures already published, particularly encompassing the active site. Though the same contacts are made between the BPL and the ligand in these structures, the overall architecture and residues that surround the site differ greatly (as can be seen in a structure based sequence alignment of BPLs active site in Figure 3a). These differences extend to eukaryotic BPLs, which contain a large N-terminal extension that shows no sequence homology to known DNA binding motifs and are approximately double the size of their prokaryotic ancestors. Truncations to the N-terminal of human BPL render an the enzyme inactive [36] indicating an underlying reliance for catalytic function on this region which is not seen for prokaryotic BPLs. Furthermore human BPL shows variation in potential ligand binding residues. SaBPL residues Ser93, Thr94, Arg120, Arg122 and Ser128 have been replaced in the human BPL enzyme by Arg480, Leu481, Lys506, Asn511 and Leu514 respectively (Figure 6). These differences can be exploited to selectively target one BPL (the pathogens) over another (the host). Due to the essential nature of this enzyme for survival and *S. aureus* being one of the most clinically prevalent pathogenic micro-organisms, our endeavor is to use the new structural information provided by SaBPL to target this protein and in the process, develop novel antibiotics to combat the growing world-wide issue of antibiotic resistance.

Figure 6

a



b

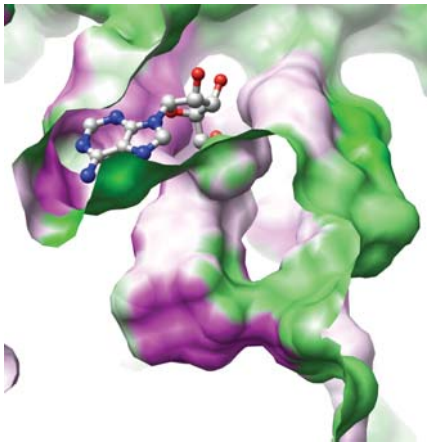


Figure 6: A structure based sequence comparison between BPLs of *S. aureus* (Biotinyl-5'-AMP-bound) *E. coli* (Biotinol-5'-AMP-bound, PDBID: 2EWN), *P. horikoshii*-K111A (Biotinyl-5'-AMP-bound, PDBID: 2DJZ) and *M. jannaschii* (Biotin-bound, PDBID: 2EJ9) highlighting structural similarities between these BPLs. (a) Shows the surface representation of the protein from above while (b) shows the surface looking inside the protein. Green shows sequence-structure differences and purple shows similarity.

Acknowledgements

The candidate thanks the staff at the Protein Crystallography Beam Line at the Australian Synchrotron for aiding the data collection of Bt-SaBPL and Apo-SaBPL.

References

1. Lane, M.D., Rominger, K.L., Young, D.L. & Lynen, F. The Enzymatic Synthesis of Holotranscarboxylase from Apotranscarboxylase and (+)-Biotin. Ii. Investigation of the Reaction Mechanism. *J Biol Chem* **239**, 2865-71 (1964).
2. Jitrapakdee, S. & Wallace, J.C. The biotin enzyme family: conserved structural motifs and domain rearrangements. *Curr Protein Pept Sci* **4**, 217-29 (2003).
3. Wakil, S.J., Stoops, J.K. & Joshi, V.C. Fatty acid synthesis and its regulation. *Annu Rev Biochem* **52**, 537-79 (1983).
4. Gravel, R.A., Lam, K.F., Mahuran, D. & Kronis, A. Purification of human liver propionyl-CoA carboxylase by carbon tetrachloride extraction and monomeric avidin affinity chromatography. *Arch Biochem Biophys* **201**, 669-73 (1980).
5. Lau, E.P., Cochran, B.C., Munson, L. & Fall, R.R. Bovine kidney 3-methylcrotonyl-CoA and propionyl-CoA carboxylases: each enzyme contains nonidentical subunits. *Proc Natl Acad Sci U S A* **76**, 214-8 (1979).
6. Samols, D. et al. Evolutionary conservation among biotin enzymes. *J. Biol. Chem.* **263**, 6461-6464 (1988).
7. Pendini, N.R. et al. Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency. *Biochim Biophys Acta* **1784**, 973-82 (2008).
8. 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**, 509-23 (2006).
9. 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).
10. Weaver, L.H., Kwon, K., Beckett, D. & Matthews, B.W. 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**, 6045-50 (2001).
11. Streaker, E.D. & Beckett, D. The biotin regulatory system: kinetic control of a transcriptional switch. *Biochemistry* **45**, 6417-25 (2006).
12. Xu, Y. & Beckett, D. 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**, 7354-60 (1994).
13. Prakash, O. & Eisenberg, M.A. Biotinyl 5'-adenylate: corepressor role in the regulation of the biotin genes of Escherichia coli K-12. *Proc Natl Acad Sci U S A* **76**, 5592-5 (1979).
14. Beckett, D. & Matthews, B.W. *Escherichia coli* repressor of biotin biosynthesis. *Methods Enzymol.* **279**, 362-376 (1997).
15. 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**, 1507-16 (2002).

16. Bagautdinov, B., Kuroishi, C., Sugahara, M. & Kunishima, N. Crystal structures of biotin protein ligase from Pyrococcus horikoshii OT3 and its complexes: structural basis of biotin activation. *J Mol Biol* **353**, 322-33 (2005).
17. Bagautdinov, B., Matsuura, Y., Bagautdinova, S. & Kunishima, N. Protein biotinylation visualized by a complex structure of biotin protein ligase with a substrate. *J Biol Chem* (2008).
18. Gupta, V. et al. Crystallization and preliminary X-ray diffraction analysis of biotin acetyl-CoA carboxylase ligase (BirA) from Mycobacterium tuberculosis. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 524-7 (2008).
19. Bhakdi, S. & Tranum-Jensen, J. Alpha-toxin of Staphylococcus aureus. *Microbiol Rev* **55**, 733-51 (1991).
20. Boyle-Vavra, S. & Daum, R.S. Community-acquired methicillin-resistant Staphylococcus aureus: the role of Panton-Valentine leukocidin. *Lab Invest* **87**, 3-9 (2007).
21. Andrews, R.K. et al. The use of snake venom toxins as tools to study platelet receptors for collagen and von Willebrand factor. *Haemostasis* **31**, 155-72 (2001).
22. Parsek, M.R. & Singh, P.K. Bacterial biofilms: an emerging link to disease pathogenesis. *Annu Rev Microbiol* **57**, 677-701 (2003).
23. Chopra, I. Antibiotic resistance in Staphylococcus aureus: concerns, causes and cures. *Expert Rev Anti Infect Ther* **1**, 45-55 (2003).
24. Enright, M.C. et al. The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). *Proc Natl Acad Sci U S A* **99**, 7687-92 (2002).
25. Foster, T.J. The Staphylococcus aureus "superbug". *J Clin Invest* **114**, 1693-6 (2004).
26. Barie, P.S. Antibiotic-resistant gram-positive cocci: implications for surgical practice. *World J Surg* **22**, 118-26 (1998).
27. Hastings, P.J., Rosenberg, S.M. & Slack, A. Antibiotic-induced lateral transfer of antibiotic resistance. *Trends Microbiol* **12**, 401-4 (2004).
28. Wang, L. & Barrett, J.F. Control and prevention of MRSA infections. *Methods Mol Biol* **391**, 209-25 (2007).
29. Tsiodras, S. et al. Linezolid resistance in a clinical isolate of Staphylococcus aureus. *Lancet* **358**, 207-8 (2001).
30. Argyrou, A., Vetting, M.W., Aladegbami, B. & Blanchard, J.S. Mycobacterium tuberculosis dihydrofolate reductase is a target for isoniazid. *Nat Struct Mol Biol* **13**, 408-13 (2006).
31. Freiberg, C. et al. Novel bacterial acetyl coenzyme A carboxylase inhibitors with antibiotic efficacy in vivo. *Antimicrob Agents Chemother* **50**, 2707-12 (2006).
32. Freiberg, C. et al. Identification and characterization of the first class of potent bacterial acetyl-CoA carboxylase inhibitors with antibacterial activity. *J Biol Chem* **279**, 26066-73 (2004).
33. Purushothaman, S., Gupta, G., Srivastava, R., Ramu, V.G. & Surolia, A. Ligand specificity of group I biotin protein ligase of Mycobacterium tuberculosis. *PLoS ONE* **3**, e2320 (2008).
34. Campbell, A., Chang, R., Barker, D. & Ketner, G. Biotin regulatory (bir) mutations of Escherichia coli. *J Bacteriol* **142**, 1025-8 (1980).
35. Chapman-Smith, A., Mulhern, T.D., Whelan, F., Cronan, J.E., Jr. & Wallace, J.C. The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity. *Protein Sci* **10**, 2608-17 (2001).
36. Campeau, E. & Gravel, R.A. 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**, 12310-6 (2001).
37. Pendini, N.R., Polyak, S.W., Booker, G.W., Wallace, J.C. & Wilce, M.C. Purification, crystallization and preliminary crystallographic analysis of biotin protein ligase from Staphylococcus aureus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 520-3 (2008).
38. Chapman-Smith, A., Turner, D.L., Cronan, J.E., Morris, T.W. & Wallace, J.C. Expression, biotinylation and purification of a biotin-domain peptide from the biotin carboxy carrier protein of *Escherichia coli* acetyl-CoA carboxylase. *Biochem. J.* **302**, 881-887 (1994).
39. Leslie, A.G. Integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **55**, 1696-702 (1999).
40. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-3 (1994).
41. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55 (1997).
42. Lovell, S.C. et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-50 (2003).
43. Delano, W.L. The PyMOL Molecular Graphics System. (DeLano Scientific, Palo Alto, 2002).
44. Gouet, P., Robert, X. & Courcelle, E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31**, 3320-3 (2003).
45. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
46. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97 (2007).
47. white, N., B., http://www.scripps.edu/~nwhite/B
48. Wilce, J.A. et al. Structure of the RTP-DNA complex and the mechanism of polar replication fork arrest. *Nat Struct Biol* **8**, 206-10 (2001).
49. Kleywegt, G.J. & Jones, T.A. Where freedom is given, liberties are taken. *Structure* **3**, 535-40 (1995).
50. Noble, M.E.M., Musacchio, A., Saraste, M., Courtneidge, S.A. & Wierenga, R.K. Crystal structure of the SH3 domain in human Fyn - Comparison of the 3-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J.* **12**, 2617-2624 (1993).
51. Cronan, J.E., Jr. Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins. *J Biol Chem* **265**, 10327-33 (1990).
52. Hayward, S. & Berendsen, H.J. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144-54 (1998).
53. Beckett, D. Biotin Sensing: Universal Influence of Biotin Status on Transcription. *Annu Rev Genet* (2007).
54. Streaker, E.D. & Beckett, D. Coupling of protein assembly and DNA binding: biotin repressor dimerization precedes biotin operator binding. *J Mol Biol* **325**, 937-48 (2003).

Chapter 5

Biotin protein ligase from *Candida albicans*: expression, purification and development of a novel assay.

This chapter presents the cloning and expression of BPL from *Candida albicans*. I was able to produce large amounts of the enzyme when expression was conducted in yeast and several rounds of crystallisation trials were attempted in 96 well trays with the enzyme alone, with the addition of biotin and/ or with ATP. Though none of the screens tried thus far produced protein crystals, the expression of large quantities of the protein allowed characterisation and development of a novel assay system to test BPL activity.

Chapter 5

Biotin Protein Ligase from *Candida albicans:* Expression, purification and development of a novel assay.

Author	Contribution	Signature
[+]Nicole R. Pendini	Purified CaBPL using bacterial and yeast expression systems. Optimised protocols for protein purification. Provided Figure 2C. Assisted in preparation of the manuscript.	
[+]Lisa M. Bailey	Conducted all enzymatic analysis of CaBPL. Produced Figure 4 and Supp Figure S2.	
Grant W. Booker	Provided intellectual discussion and assisted in manuscript preparation.	
Matthew C. Wilce	Provided intellectual discussion and assisted in manuscript preparation.	
*John C. Wallace	Provided intellectual discussion and assisted in manuscript preparation.	
*Steven W. Polyak	Primary role in project management and manuscript preparation. Performed all recombinant DNA cloning including construction of expression vectors. Provided Figures 1, 2A&B, 3 and Supp Figure S1.	

[+] = Equal first authorship

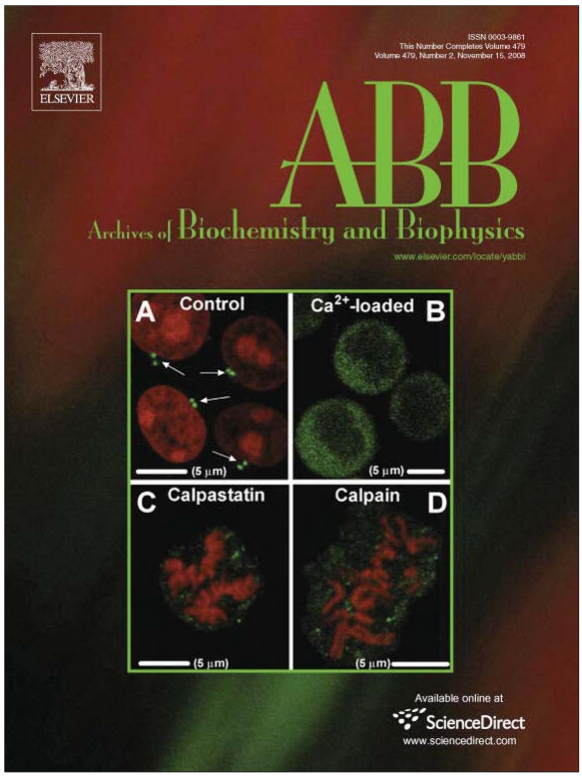
This work was supported by funding from a commercial partner awarded to JCW and SWP.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Sep 15, 2008

This is a License Agreement between Nicole R Pendini ("You") and Elsevier ("Elsevier"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Nicole R Pendini
Customer address	biochemistry, bld 13D, wellington rd Victoria, other 3800
License Number	2030480538982
License date	Sep 15, 2008
Licensed content publisher	Elsevier
Licensed content publication	Archives of Biochemistry and Biophysics
Licensed content title	Biotin Protein Ligase from *Candida albicans*: Expression, purification and development of a novel assay
Licensed content author	Nicole R. Pendini, Lisa M. Bailey, Grant W. Booker, Matthew C. J. Wilce, John C. Wallace and Steven W. Polyak
Licensed content date	11 September 2008
Volume number	n/a
Issue number	n/a
Pages	1
Type of Use	Thesis / Dissertation
Portion	Full article
Format	Electronic
You are an author of the Elsevier article	Yes
Are you translating?	No
Purchase order number	
Expected publication date	Oct 2008



Archives of Biochemistry and Biophysics 479 (2008) 163–169

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Biotin protein ligase from *Candida albicans*: Expression, purification and development of a novel assay

Nicole R. Pendini [a,1], Lisa M. Bailey [a,1,2], Grant W. Booker [a], Matthew C.J. Wilce [b], John C. Wallace [a], Steven W. Polyak [a,*]

[a] School of Molecular and Biomedical Science, University of Adelaide, Molecular Lifesciences Building, North Tce, Adelaide, SA 5005, Australia
[b] Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Vic., Australia

ARTICLE INFO

Article history:
Received 25 July 2008
and in revised form 29 August 2008
Available online 11 September 2008

Keywords:
Biotin protein ligase
Candida albicans
Fungicide, drug discovery
Protein structure and function
Assay development

ABSTRACT

Biotin protein ligase (BPL) is an essential enzyme responsible for the activation of biotin-dependent enzymes through the covalent attachment of biotin. In yeast, disruption of BPL affects important metabolic pathways such as fatty acid biosynthesis and gluconeogenesis. This makes BPL an attractive drug target for new antifungal agents. Here we report the cloning, recombinant expression and purification of BPL from the fungal pathogen *Candida albicans*. The biotin domains of acetyl CoA carboxylase and pyruvate carboxylase were also cloned and characterised as substrates for BPL. A novel assay was established thereby allowing examination of the enzyme's properties. These findings will facilitate future structural studies as well as screening efforts to identify potential inhibitors.

Systemic fungal infections cause life-threatening invasive disease especially for immunocompromised and/or debilitated patients [1,2]. Whilst there are many different fungal pathogens *Candida* spp. remains one of the most important in the clinic. *Candida* is the fourth most common cause of nosocomial bloodstream infections in the United States [3]. The treatment failure rate for invasive candidiasis remains high at 15–40% [4–6]. *Candida* species are also a major cause of superficial diseases, particularly vulvovaginal infections in women of child-bearing age and oral infections in neonates and the elderly [7,8]. These infections are a major cost burden on healthcare systems due to prolonged treatment in intensive care units, mechanical ventilation and haemodialysis. Adding to these pressures is the increased incidence of *Candida* strains resistant to current therapeutics [3,9,10]. In order to develop improved approaches to antifungal therapy, new targets, novel compounds and innovative delivery systems are required [1,11]. The identification of key enzymes that represent novel drug targets in pathogenic fungi and small molecule inhibitors are basic necessities of the drug discovery pipeline.

Biotin protein ligase (BPL) is an essential metabolic enzyme (reviewed [12]) and, therefore, represents a novel antifungal drug target. Whilst the BPL from *Candida albicans* has not previously been characterised, the model fungal organism *Saccharomyces cerevisiae* has provided valuable insight. A comparison of the genomic sequences from *S. cerevisiae* [13] and *C. albicans* SC5314 [14] shows both contain a single *bpl* gene (39% identical at the protein level). Disruption of one *bpl* allele in diploid *S. cerevisiae* cells resulted in only half of the spores being viable after sporulation and none of these viable cells were BPL negative [15]. This result emphasises the essential function of BPL. The enzyme is responsible for the *in vivo* biotinylation of all six biotin-dependent enzymes in *S. cerevisiae* [15]: two isoforms of acetyl CoA carboxylase (ACC), two isoforms of pyruvate carboxylase (PC), urea amidolyase and the recently identified tRNA binding protein Arc1p. Cytoplasmic ACC is the rate-limiting step in lipogenesis where it supplies malonyl CoA for long chain fatty acid elongation and membrane biogenesis [16]. This is an essential enzyme as disruption of the *acc1* gene is lethal to spores, even when grown with fatty acid supplementation [16]. The mitochondrial form of ACC, HFA, also participates in fatty acid synthesis but has been implicated specifically in the synthesis of the cofactor lipoic acid required for oxidative decarboxylation of glycine and several alpha-keto acids including pyruvate [17]. PC is also essential in yeast where it plays a key anaplerotic role during gluconeogenesis [18–20]. Double knockout strains have been generated but these can only be maintained for a limited time even with supplementation of the growth media with aspartate [18–20]. The fifth known biotin enzyme in yeast is urea amidolyase which catalyses the degradation of urea. Recently, the tRNA binding protein, Arc1p, was also shown to incorporate biotin although

* Corresponding author. Fax +61 8 8303 4362.
E-mail address: steven.polyak@adelaide.edu.au (S.W. Polyak).
[1] Equal first authorship.
[2] Present address: The Royal Institution of Great Britain, 21 Albemarle Street, London W1S 4BS UK.

0003-9861/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2008.08.021

biotin is not required for the function of this non-essential protein [21].

Here we report isolation of the coding region of the BPL from *C. albicans* (CaBPL) and procedures for recombinant expression and purification. Biotin domains from ACC and PC were also characterised as substrates for CaBPL. These were employed to establish a new assay for CaBPL.

Materials and methods

Materials

Europium-labelled streptavidin was from Perkin-Elmer. All other materials were as previously described in [22]. All reagents were of analytical grade or higher.

DNA manipulations

Cloning of C. albicans BPL

Genomic DNA from *C. albicans* CBF562 was purified as previously described [23]. The gene for CaBPL was obtained by genomic PCR. Sequence data for oligonucleotide design were obtained from GenBank (GI:46441220). The sequences for the oligonucleotides employed in this study are shown in Table 1. The gene was obtained as three separate fragments, thus facilitating modification of the DNA sequence.

A fragment spanning the 5′ half of the gene, encompassing nucleotides 1–1016, was obtained using PCR overlap extension. Here two smaller genomic PCR products were initially obtained either with oligonucleotides C1/32 and C2/39, or with C3/44 and C4/44 upon genomic DNA. These two products were purified and included as a template for PCR with oligonucleotides C1/32 and C4/44. This produced a 1027 base pair fragment which was cloned into pGEM-T Easy, producing pGEM(C1/C4)Ser. Through this approach a BspH1 restriction site was introduced at the initiation codon and a unique Nar1 restriction site engineered at nucleotide 411. In addition the CUG codons at nucleotides 402 and 432 were altered to the serine codons AGU and AGC, respectively.

A second fragment was obtained that spanned nucleotides 567–1016 using genomic PCR with oligonucleotides C5/20 and C4/44. The 449 base pair fragment was subcloned into pGEM-T Easy generating pGEM(C5/C4). An Acc1 restriction site at nucleotide 859 and the CUG codon at position 865 were simultaneously modified using oligonucleotides C10/45 and C11/45 with the QuikChange®

Table 1
Oligonucleotides employed in this study

C1/32	5′^{1}ATCTACTCATGA^{1}ATGTTTTAGTATATTCTGGC^{20}
C2/39	5′^{425}GACGGCTCTGGCGCCAGTACGACTT**TCG**TACTTGAAACC^{387}
C3/44	5′^{405}CGTACTGGCGCCAGAGCCGTCAAATTG**AGC**GTCAATACAGCTGC^{449}
C4/44	5′^{1016}TCCGGTATACCGGC**GCT**AGTGAGAAAGTTGATATACTTGACTGG^{973}
C5/20	5′^{567}GTGGATTTAGAGAAGGCTGC^{586}
C6/22	5′^{943}ACTAGTGAGTATGTTGGTAGTC^{964}
C7/65	5′GAGCTCGGTACCTAATGATGATGATGATGATGATGATGACCGGT^{1992} CTTCTTATATACTAAACC^{1975}
C10/45	5′^{843}GGATAAAGTTAGGGATGT**G**TAC**AGC**ATTTTAACCAGCAAGTTAGG^{887}
C11/45	5′^{887}CCTAACTTGCTGGTTAAAAT**GCT**GTA**C**ACATCCCTAACTTTATCC^{843}
C12/33	5′ATCTACGGATCC^{2149}AAGGAAGAGGCATCTGCCACT^{2169}
C13/38	5′ATCTACGGATCC^{2197}TTATTAGAAGTTGAAAATGATCCAAC^{2222}
C14/37	5′ATCTACGAATTCATCA^{2433}ATCGTCCAATGCCAAAATGGC^{2413}
C15/40	5′^{700}GTGGTTGACACACTTCGAGCATACGATCACAACAAAAAGG^{739}
C16/40	5′^{739}CCTTTTTGTTGTGATCGTATGCTCGAAGTGTGTCAACCAC^{700}
C17/34	5′ATCTACGGATCC^{2668}ATGAGATCAGTTTCCGTTGAAG^{2689}
C18/38	5′ATCTACGGATCC^{2725}AAAGCTTCAGCATCAAATGAAGTTGC^{2750}
C19/43	5′ATCTACGAATTC^{2946}ATCAATGAATACTAGTAATCAAATCATTAGC^{2920}
C20/28	5′GGATCC^{2602}ATGGCTGTTGGTGATGTTTCGG^{2623}

Endonuclease restriction sites are underlined and mutagenic sequences are shown in bold. Numbers in superscript designate position in the target gene sequence.

mutagenesis protocol (Stratagene), producing pGEM(C5/C4)Ser. A polymorphism different from the published sequence in the species used here was detected by DNA sequencing. Nucleotide 719 was found to be a T base thus introducing another Acc1 site into the gene and changing the codon to an alanine (published sequence contains a valine at the corresponding position). The nucleotide at this position was mutated to C with oligonucleotides C15/40 and C16/40, producing pGEM(C5/C4)Ser-Acc.

The third of the overlapping fragments encompassed the entire 3′ half of the C*abpl* gene, from nucleotides 943 to the termination codon at 1993. This was generated by genomic PCR with oligonucleotides C6/22 and C7/65 and cloned into pGEM-T Easy, producing pGEM(C6/C7). The C7/65 primer fused the coding sequence for a glycine-threonine-(histidine)$_9$ extension onto the C-terminus of the expressed gene product. Restriction sites for endonucleases Kpn1 and Sac1 were also engineered onto the end of the gene to facilitate subcloning.

The 5′ and 3′ halves of the CaBPL gene were fused together by digesting pGEM(C1/C4)Ser with Spe1 and Sac1 and ligation with similarly treated fragment from pGEM(C6/C7). The sequence between nucleotides 592 and 1007 in the resulting vector was replaced with the modified sequence from pGEM(C5/C4)Ser-Acc using the Acc1 sites at these positions. The final construct, pGEM(caBPL-His$_9$) contained the full length gene with all the desired modifications. For the recombinant expression of BPL for *in vivo* complementation assays, the BspH1/Kpn1 fragment from pGEM(caBPL-His$_9$) was cloned into Nco1/Kpn1 treated vector pARA13 [24]. For high-level expression of BPL for purification from *E. coli*, the BspH1/Sac1 fragment from pGEM(caBPL-His$_9$) was cloned into Nco1/Sac1 treated pET-16b (Novogen). The Xba1/Sal1 fragment from pET(CaBPL-His$_9$) was subsequently subcloned into Xba1/Xho1 treated pVT100 u vector [25] to generate pVT(CaBPL-His$_9$) for yeast cell expression.

Cloning biotin domains

A series of plasmids suitable for expressing peptides encompassing the predicted biotin domain of *C. albicans* ACC and PC were generated. Here three polypeptides from PC and two from ACC were investigated. The PC polypeptides corresponded to the C-terminal −74 (residues 909–982), −93 (residues 890–982) and −115 (residues 868–982) amino acids of PC respectively, with the biotinylated lysine at position 948. The ACC polypeptides corresponded to two internal sequences of −94 (residues 717–811) and −78 (733–811) amino acids, with the biotinylated lysine at position 775. The DNA encoding each peptide was amplified using PCR upon genomic DNA. Sequence data for oligonucleotide design were obtained from GenBank (PC GI:46442935; ACC GI:46440402). To generate each construct the following combinations of oligonucleotides were employed in the PCR; C12/33 and C14/37 for the 78 amino acid polypeptide (caACC-78), C13/38 and C14/37 for the 94 amino acid polypeptide (caACC-94), C18/38 and C19/43 for the 74 amino acid polypeptide (caPC-74), C17/34 and C19/43 for the 93 amino acid polypeptide (caPC-93) and C20/28 and C19/43 for the 115 amino acid polypeptide (caPC-115). The above oligonucleotides introduced a BamH1 restriction site at the 5′ end of the PCR product and an EcoR1 site at the 3′ end. The PCR products were digested with BamH1 and EcoR1 restriction endonucleases and ligated into similarly treated pGEX-4T-2 plasmid (Amersham-Biosciences).

Recombinant protein expression and purification

CaBPL-H$_9$. The yeast expression vector pVT(CaBPL-H$_9$) was transformed into *S. cerevisiae* w303 [26]. Cell lysate containing recombinant CaBPL-H9 was prepared as described previously [27] with the following modifications: spheroplasts were resuspended in Yeast Buster® solution (Novagen) at 5 ml/g cell pellet. This buffer was

supplemented with 1 mM tris(hydroxypropyl)phospine, 35 mM PMSF and Ni-NTA Protease Inhibitor Cocktail (Merck). The spheroplasts were disrupted with gentle agitation for 1 h at room temperature. Cleared, filtered lysates were then fractionated by NiNTA chromatography, as previously described [27]. Fractions containing BPL activity were pooled and exchanged into buffer A (50 mM $NaPO_4$ pH 6.0, 5% glycerol, 1 mM EDTA, 1 mM DTT). To further purify CaBPL, cation exchange chromatography was performed. A 20 mL S-sepharose column (GE Healthcare) was equilibrated in buffer A prior to application of the NiNTA purified sample. The mixture was then fractionated with a 0–500 mM NaCl gradient over 10 column volumes. CaBPL eluted with 330 mM NaCl, as determined by NiNTA blot and BPL activity assay. This material was pooled and exchanged into storage buffer (50 mM Tris–HCl pH 8.0, 0.5 mM EDTA pH 8.0, 5% glycerol, 1 mM DTT). SDS–PAGE analysis revealed the preparations were >95% pure. Samples were stored at −80 °C.

CaPC115. CaPC115 was expressed as a GST fusion protein permitting high-level expression in *E. coli* and rapid purification by affinity chromatography. Protein over-expression was performed as previously described [28] with the following modifications; cells were grown in 2YT media and disrupted by two passages through a French Press (42,000–60,000 kPa) prior to glutathione affinity chromatography. The filtered lysate was passed over a 1 ml GSTTrap column (GE Healthcare) continuously overnight at 4 °C. Unbound material was removed by washing with 10 column volumes of PBS containing 1 mM DTT. The column was equilibrated in 5 volumes of thrombin digestion buffer (20 mM Tris–HCl pH 8.5, 150 mM NaCl, 2.5 mM $CaCl_2$) before addition of 7.5 U of biotinylated thrombin (Novagen). The GST fusions were cleaved overnight at RT before the cleaved biotin domains were washed off the column in 5 volumes of thrombin digestion buffer. To isolate the apo-biotin domain, biotinylated thrombin and holo-biotin domain were simultaneously removed from solution using Streptavidin-Sepharose High Performance (Amersham Biosciences) in a pull-down reaction, following manufacturers' instructions. Apo-CaPC-115 in the supernatant was collected, dialysed against 2 mM ammonium acetate pH 7.4 and stored lyophilised at −20 °C. The purification was monitored by SDS–PAGE and streptavidin blot to ensure complete removal of biotinylated protein [Supp Fig. 1].

BPL assay. Two methods were employed for assaying CaBPL activity. The incorporation of ^3H biotin onto an acceptor substrate was performed as previously described [29,30]. A 96-well assay, using Europium-labelled streptavidin and time-resolved fluorescence, was also developed. Here, Apo-CaPC115 was diluted in Tris-buffered saline (TBS, 20 mM Tris–HCl, pH 7.2, 150 mM NaCl) and 100 ng/well added to white Lumitrac-600 96-well plates (Greiner Bio-one). Varying amounts (0–100 ng) of holo-CaPC115 were also included on each plate, to serve as standards for product formation. After coating overnight at 4 °C, plates were blocked in TBS with 1% BSA for 1 h, followed by five washes in TBS, 0.1% Tween. For the biotinylation assay, 50 μL of master mix containing 50 mM Tris–HCl pH 8.0, 5 mM biotin, 50 μM ATP, 5.5 mM $MgCl_2$, 0.1 mg/mL BSA and 0.1 mM DTT was added to each well. The reaction was initiated by the addition of CaBPL to a final concentration of 3.4 nM, and terminated by the addition of EDTA to 55 mM. Reactions were performed for 15 min at 37 °C after which the plate was washed five times in TBS, 0.1% Tween. Europium-labelled Streptavidin (Perkin-Elmer) was diluted to 0.1 μg/ml in TBS containing 0.1% Tween and 100 μM 1,3-diaminopropane-2-ol-*N,N′*-tetraacetic acid (DPTA; Sigma, St. Louis, MO, USA). The Streptavidin probe (50 μL per well) was incubated for 1 h at 37 °C, followed by washing five times in TBS 0.1% Tween containing 100 μM DPTA followed by three washes in sterile water. Enhancement Solution (Perkin-Elmer) (50 μL) was added per well and incubated for 10 min before reading the plate using time-resolved fluorescence using 340-nm excitation and 612-nm emission filters with a PolarStar Galaxy microplate reader (BMG Labtech, Vic, Australia). One well containing 0.8 pmol (10 ng) of holo biotin domain was routinely employed to adjust the gain, thus ensuring a linear response for 0–3.2 pmol (0–40 ng) of product.

Results and discussion

Cloning of CaBPL-H_9 and recombinant production in E. coli

The *C. albicans* BPL (CaBPL) gene contains 4 rare CUG codons, which are read as serine in Candida but code for leucine in the universal genetic code. Therefore a cloning strategy was devised that permitted modification of all four codons to universal serine codons, necessary for recombinant expression of the BPL in a bacterial host system. Additionally, CaBPL was fused with 9 histidine residues on the C-terminus to facilitate purification by immobilised metal ion affinity chromatography. The gene was amplified by genomic PCR as three overlapping fragments and then assembled into a single coding region. The integrity of the modified gene was verified in a complementation assay using a conditionally lethal birA$^-$ strain of *E. coli* [31] [Fig. 1].

In order to characterise this protein further, the modified coding region was also introduced into the pET16 b expression vector for recombinant enzyme production in *E. coli*. The enzyme was expressed in *E. coli* Rosetta cells grown at 30 °C in an attempt to minimise inclusion body formation. Protein expression was induced for 3 h with IPTG before the bacterial cells were harvested and lysed. The cell lysate was loaded onto a Nickel-NTA chelating column in the presence of 10 mM imidazole, washed with 50 mM imidazole and CaBPL finally eluted with 100–500 mM imidazole. Fractions containing purified CaBPL were immediately dialysed against storage buffer (50 mM Tris–HCl pH 8.0, 5% glycerol, 0.1 mM EDTA, 1 mM DTT) to remove the imidazole and prevent precipitation. The enzyme preparation was measured with an *in vitro* biotinylation assay and shown to have a specific activity of 49 nmol/min/mg. This is a similar specific activity to that reported for yeast BPL after nickel-chelating chromatography [29] but lower than *E. coli* BPL purified by two ion exchange steps [32], thus highlighting the requirement for a further purification step. SDS–PAGE and Western blot probed with a Ni-NTA alkaline phosphatase conjugate revealed a major purified product that migrated to 73 kDa, the expected molecular mass for CaBPL-H_9 [Fig. 2A and B]. N-terminal sequencing revealed this species to con-

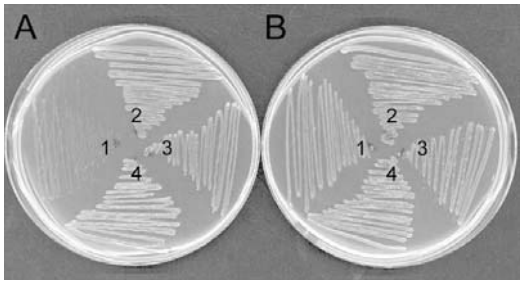


Fig. 1. Complementation assay. The conditional lethal *birA*85 *E. coli* strain BM4062 cell was transformed with plasmids (1) pAra13 [24], (2) pAra(yBPL) [29], (3) pAra(CaBPL) (this study) and (4) pAra(hBPL) [40] that express no BPL, or BPL from *S. cerevisiae*, *C. albicans* or *Homo sapiens* respectively. Strains were grown at either (A) the restrictive temperature of 42 °C or (B) the permissive temperature of 30 °C. Growth at 42 °C demonstrates recombinant expression of a functional BPL.

tain the first six amino acids for the protein (i.e. MNVLVY). In addition, a 48 kDa protein that reacted with the Ni-NTA conjugate co-purified with full length CaBPL-H_9. N-terminal sequencing detected the motif MRDCLK, suggesting this was a C-terminal fragment of CaBPL resulting from proteolysis between F248 and M249. Previous analysis of yeast BPL production in *E. coli* reported an analogous proteolytic product and is indicative of a protease sensitive linker that is proposed to connect an N-terminal domain to the catalytic C-terminal region [29]. Due to difficulties in purifying the cleaved protein from intact CaBPL, an alternative expression system was investigated.

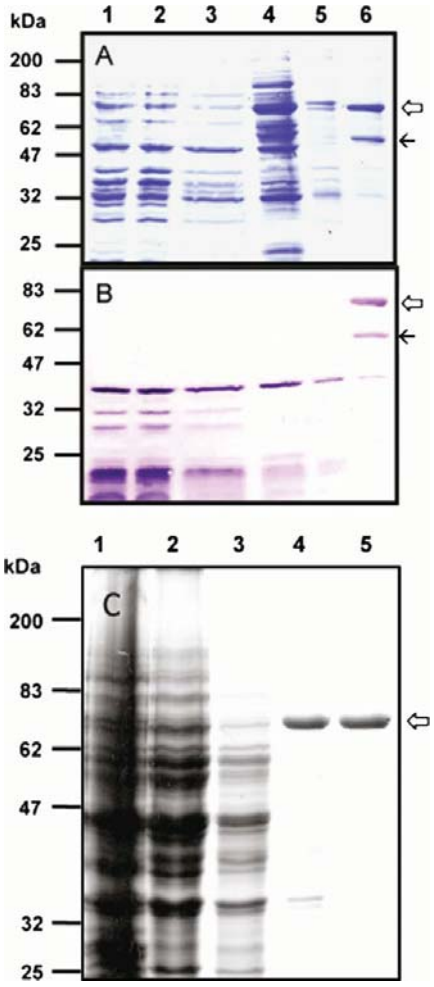


Fig. 2. Preparation of CaBPL. Expression of recombinant CaBPL was performed in *E. coli* Rosetta cells (A and B) and yeast (C), as described in Materials and methods. Purification of the bacterially derived protein using NiNTA chromatography was monitored by (A) SDS–PAGE and (B) Western blot probed with NiNTA—alkaline phosphatase. Samples analysed were (1) 50 μg of the soluble fraction of bacterial cell lysates, (2) 50 μg unbound material, (3) 20 μg of the wash fraction, (4) 50 μg of protein eluted with 100 mM imidazole, (5) 10 μg of protein eluted with 250 mM imidazole and (6) 10 μg of protein eluted with 500 mM imidazole. (C) Preparation of yeast-derived CaBPL was monitored by SDS–PAGE. Samples analysed were (1) 50 μg of the cell lysate obtained with Yeast Buster Reagent, (2) 50 μg of cell lysate after sonication (3) 20 μg unbound material, (4) 10 μg of protein after NiNTA chromatography and (5) 10 μg of protein after cation exchange. The migration of molecular mass standards is shown to the left of the gel. Bands corresponding to full length CaBPL (⇦) and the truncated product formed by *in vivo* proteolysis (←) are shown.

Recombinant production of CaBPL-H_9 in yeast

To improve production of CaBPL a yeast expression system was also investigated. It was proposed that expression in this host would overcome issues with proteolysis and solubility. The coding region of CaBPL was introduced into the yeast vector pVT100 u for constitutive protein expression from the strong alcohol dehydrogenase promoter [25]. Analysis of CaBPL expression by Western-blot revealed optimal protein production after 48 h growth at 30 °C. CaBPL was then isolated from whole cell lysates by NiNTA chromatography. SDS–PAGE and NiNTA blot analysis showed the prominent species obtained from the purification migrated to a molecular mass expected for CaBPL (ie 73 kDa) [Fig. 2C]. Importantly, the 48 kDa proteolytic product observed during expression in *E. coli* was not present. To remove the remaining contaminating proteins cation exchange chromatography was performed using S-Sepharose resin. Fractionation of the sample with a 0–500 mM NaCl gradient over 10 column volumes was sufficient to resolve the contaminating proteins from CaBPL [Fig. 2C]. With this two-chromatography step procedure we were able to purify CaBPL to homogeneity with yields of up to 60 mg per litre culture, a substantial improvement over the bacterial system. We observed that the protein could be concentrated to 13 mg/ml and stored for up to 7 days at 4 °C at this concentration without precipitation.

Characterisation of biotin domains

An additional requirement for this work was to identify a suitable substrate to assay CaBPL. A comparison of the genomes of *C. albicans* and *S. cerevisiae* revealed differences in the number of biotin-dependent enzymes between the two organisms. Like *S. cerevisiae*, the *C. albicans* genome contains homologues for urea amidolyase (66% identity) and Arc1p (47% identity). In contrast to *S. cerevisiae*, *C. albicans* has only single genes for ACC (69% identity vs ACC, 53% identity vs HFA) and PC (76% identity vs PC1, 76% identity vs PC2). Given the important metabolic roles played by both ACC and PC we decided to focus upon these two enzymes and define the biotin domains. A series of polypeptides encompassing the biotin attachment site were assayed using an *in vivo* biotinylation assay in *E. coli* BL21. This approach facilitated rapid analysis of the poly peptides as BPL substrates [30,33]. Each polypeptide was expressed as a GST fusion to facilitate quantitative analysis of the polypeptides and downstream purification. To determine if polypeptides were functional BPL substrates, whole cell lysates containing the over-expressed GST-fusions were assessed by Western blot analysis. Here, blots were probed with both Streptavidin-HRP, to detect biotinylated products, and an anti-GST antibody to monitor protein expression. The anti-GST Western blots showed all proteins were readily overexpressed in *E. coli*. Streptavidin blot analysis showed all the three polypeptides from PC were biotinylated, implying these do indeed function as BPL substrates [Fig. 3]. Curiously, the peptides derived from ACC were poorly biotinylated. As biotin domains must adopt an appropriate structure for substrate recognition by BPL, a failure to be biotinylated suggests impairment with protein folding [34,35]. Consequently, all further analysis was performed using the pyruvate carboxylase biotin domains. Furthermore, our previous studies on biotin domains from yeast PC have shown that the longer polypeptide is more stable during purification. Therefore CaPC115 was selected in all further work.

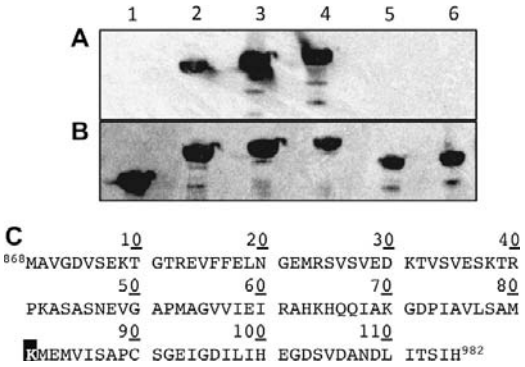


Fig. 3. *In vivo* biotinylation assay on the biotin domains of CaPC and CaACC. (A) Proteins biotinylated *in vivo* by *E. coli* BPL were visualised by Streptavidin-blot. (B) The same Western blots were also probed with an anti-GST antibody, which served as a loading control. Samples analysed were (1) GST, (2) GST-CaPC-74, (3) GST-CaPC-93, (4) GST-CaPC-115, (5) GST-CaACC-94 and (6) GST-CaACC-94. Polypeptides that react with Streptavidin were determined to be BPL substrates. (C) The protein sequence of CaPC-115 is shown. The lysine targeted for biotinylation is in black box.

Assay design

Previously published methods for measuring *in vitro* biotinylation have involved the laborious, but reliable, process of measuring incorporation of radiolabelled biotin onto discrete biotin domains [29,30,36]. We recently reported an alternative BPL assay that employs a peptide substrate for *E. coli* BPL and fluorescence polarisation [22]. Whilst this assay is amenable to high-throughput applications, the peptide's high specificity for the *E. coli* enzyme limits its use for other BPLs. Recently a peptide substrate for yeast BPL was reported [37]. However, this peptide is a poor substrate (K_M = 130 μM) relative to a biotin domain (K_M = 1 μM, [30]) thereby limiting its application in kinetic analysis where saturating concentrations of substrates are required. Here we report a modified assay system that is amenable to high throughput applications without the hazards associated with handling radioactivity. Furthermore, this assay employs a biotin domain as the BPL substrate, so can readily be adapted for use with BPLs from a wide range of species. Initially apo CaPC-115 substrate was adsorbed onto the surface of a 96-well microtitre plate. An assay solution containing buffer, biotin and MgATP is added to each well and the reaction initiated by addition of enzyme. After termination of the reaction with EDTA, the amount of immobilised, holo CaPC-115 product can be detected and quantitated using Europium-labelled streptavidin and time resolved fluorescence.

Establishment of the CaBPL assay

Conditions for immobilising CaPC-115 were first investigated. The biotin domain was readily absorbed onto microtitre plates with various buffer solutions (TBS, PBS, 50 mM Tris–HCl) at pHs compatible with BPL assays (ie pH 7.5–8.0). Following blocking with BSA, the microtitre plates could be stored at −20 °C for up to 1 week. Using 0.8 pmol (10 ng) of holo CaPC-115 as a standard, the intra-assay variation was assessed (n = 25). A coefficient of variance of 13% suggested good reproducibility in the coupling protocol across the microtitre plate. Standard curves were then produced to determine sensitivity and dynamic range. The previously characterised biotin domain holo yPC-104 was employed alongside holo-CaPC-115 [30,33]. The minimal detection limit was set as a function of the mean for the zero enzyme control plus 3× the standard deviation. Here we routinely calculated 5–10 fmol holo biotin domain as the minimal quantity that could be accurately measured [Fig. 4A]. This is comparable in sensitivity with other published assays [36,38,39]. A linear response was routinely observed between 0 and 3.2 pmol (40 ng) of the biotinylated standards.

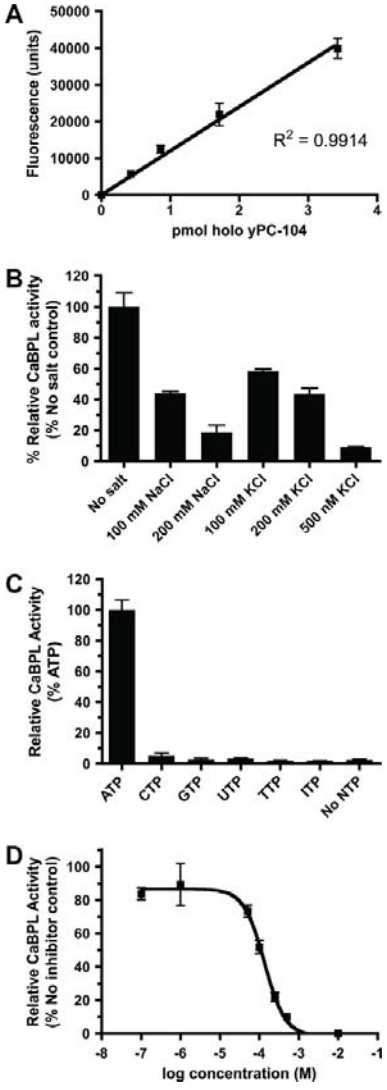


Fig. 4. Characterisation of CaBPL activity. (A) Calibration curve of product formation. A standard curve of holo-yPC104 [30] vs fluorescence was produced by adsorbing varying quantities of holo-yPC-104 onto white 96-well plates then quantifying with Europium-labelled streptavidin and time-resolved fluorescence. (B) Effect of monovalent ions on CaBPL activity. Enzyme activity was measured with the inclusion of varying concentrations of NaCl or KCl. Percentage activity is based on 100% activity with MgATP alone. (C) Activity of CaBPL with various nucleotide triphosphates. Enzyme activity was measured with the inclusion of 3 mM of selected nucleotide triphosphates in the reaction medium. Percentage activity is based on 100% activity with MgATP. (D) Inhibition of CaBPL by pyrophosphate. The activity of CaBPL was measured in the presence of varying concentrations of pyrophosphate. Percentage activity is based on 100% activity with no inhibitor.

Table 2
Activity of CaBPL with 5.5 mM of various metal ions

Metal ion	Relative BPL activity
$MgCl_2$	100%
$CaCl_2$	35 ± 0.1
RbCl	5 ± 1.2
CsCl	7 ± 1.9
$MnCl_2$	15 ± 0.9
$CoCl_2$	12 ± 3.0
$ZnCl_2$	5 ± 1.4
No metal	8 ± 0.5

Characterisation of CaBPL

The CaBPL reaction was subsequently optimised with an appropriate enzyme concentration and time point to achieve between 0.8 and 3.2 pmol (10–40 ng) product formation (ie <40% of the total reaction). The assay was routinely performed at 37 °C for 15 min, at which point the reaction was terminated by the addition of 55 mM EDTA (ie a 10-fold stoichiometric excess over Mg^{2+}). The reaction was further optimised through analysis of assay variables such as ionic strength and/or monovalent ions, as well as the requirement for metal ions and nucleotide triphosphate. The reaction was performed in several buffers to test the effect of pH on enzyme activity. Here sodium acetate buffer (pH 5.0–6.0), Tris–HCl buffer (pH 7.0–8.0) or sodium carbonate buffer (pH 9.0–10.0) was included. CaBPL was active across a broad pH range (6–10) but displayed optimal activity at pH 8.0. Addition of 100 mM NaCl and KCl inhibited the reaction to 44 ± 1.3% and 58 ± 1.4% respectively. CaBPL activity was reduced to <20% with 200 mM NaCl and 500 mM KCl [Fig. 4B].

A number of reports on the characterisation of BPLs from a wide range of species have shown that BPLs do differ in their requirements for metal ions and nucleotide sources (summarised in [29]). Therefore, $MgCl_2$ was replaced in the reaction by a series of mono- or divalent metal ions [Table 2]. As expected, EDTA was effective in terminating the reaction. The inclusion of $MgCl_2$ produced optimal CaBPL activity and was clearly the most preferred source of metal ions. $CaCl_2$ (35%), $MnCl_2$ (15%) and $CoCl_2$ (12%) could be substituted to varying degrees. All other ions tested gave <10% enzyme activity [Table 2]. Interestingly, in the absence of any metal ion residual activity was measured (8%) suggesting co-purification of $MgCl_2$ (or other divalent cation) in the enzyme preparation. Replacing ATP in the assay with CTP, GTP, UTP, TTP or ITP next addressed nucleotide triphosphate specificity. ATP was optimal, with all other NTPs yielding <5% activity [Fig. 4C].

Finally we investigated the assay with inhibition studies. As the first partial reaction is accompanied by the release of pyrophosphate in the reaction mechanism, product inhibition studies with pyrophosphate were performed. The compound had an IC_{50} of 100 μM [Fig. 4D], which is similar to results obtained with *S. cerevisiae* and *E. coli* BPL [22,29]. Furthermore DMSO did not affect CaBPL activity. A comparison of product formation in the presence of 0 or 2% DMSO showed no differences at either 5 min ($p = 0.26$) or 15 min ($p = 0.69$). This is advantageous when screening libraries of small molecules often solubilized in this solvent. Taken together, these data demonstrate that CaBPL has similar properties to those of BPL from *S. cerevisiae*.

Conclusions

We have established a protocol for the production of highly purified recombinant BPL from the pathogenic fungi *C. albicans*. The protocol employs two chromatography steps to obtain protein that is >95% pure. The use of yeast as a heterologous expression system bypassed issues with proteolysis that were observed with expression in *E. coli*. Interestingly, the protease-susceptible bond in CaBPL is analogous to a site characterised in yeast BPL [29]. It has been proposed that this site resides within a protease-sensitive inter-domain linker connecting an N-terminal domain with the C-terminal catalytic domain. Importantly, the N-terminal domain is essential for yeast BPL activity in a mechanism that is yet to be deciphered. Unlike the catalytic core of all BPLs, the N-terminal regions vary greatly in size and function [12]. An absolute dependence upon the N-terminal domain for catalysis appears to be an important feature of yeast and fungal BPLs. Given the similarities between BPLs from *S. cerevisiae* and *C. albicans* at the primary structure level (39%), predicted domain structure and kinetic properties, we propose a conserved catalytic mechanism. Ultimately a high resolution X-ray crystal structure will be invaluable in delineating the role of the N-terminal domain in catalysis. Hitherto, production of CaBPL has restricted structure determination efforts. The methodology we report here is suitable for producing sufficient quantities of purified enzyme to now support crystallisation efforts.

Acknowledgments

This work was supported the Biochemistry Discipline, School of Molecular and Biomedical Sciences, University of Adelaide. S.W.P., G.W.B. and J.C.W. were recipients of a Commercial Development Initiative Award from BioinnovationSA. M.C.J.W. acknowledges support from the National Health and Medical Research Council and the Australian Research Council.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2008.08.021.

References

[1] T.J. Walsh, M.A. Viviani, E. Arathoon, C. Chiou, M. Ghannoum, A.H. Groll, F.C. Odds, Med. Mycol. 38 (2000) 335–347.
[2] L. Ajello, R. Hay, in: T. Wilson (Ed.), Microbiology and Microbial Infections, London, UK, 1998.
[3] R.C. Moellering Jr., J.R. Graybill, J.E. McGowan Jr., L. Corey, Am. Soc. Microbiology. Am. J. Infect. Control 35 (2007) S1–S23.
[4] P.G. Pappas, J.H. Rex, J. Lee, R.J. Hamill, R.A. Larsen, W. Powderly, C.A. Kauffman, N. Hyslop, J.E. Mangino, S. Chapman, H.W. Horowitz, J.E. Edwards, W.E. Dismukes, Clin. Infect. Dis. 37 (2003) 634–643.
[5] P. Sandven, Rev. Iberoam, Micol. 17 (2000) 73–81.
[6] M.A. Slavin, J. Antimicrob. Chemother. 49 (2002) 3–6.
[7] R.A. Calderone, *Candida* and Candidiasis, ASM Press, Washington, DC, 2002.
[8] F.C. Odds, Bailliere Tindall, London. (1988).
[9] S.R. Norrby, C.E. Nord, R. Finch, Lancet Infect. Dis. 5 (2005) 115–119.
[10] G.H. Talbot, J. Bradley, J.E. Edwards Jr., D. Gilbert, M. Scheld, J.G. Bartlett, Clin. Infect. Dis. 42 (2006) 657–668.
[11] A.C. Pasqualotto, D.W. Denning, J. Antimicrob. Chemother. 61 (Suppl 1) (2008) i19–30.
[12] N.R. Pendini, L.M. Bailey, G.W. Booker, M.C. Wilce, J.C. Wallace, S.W. Polyak, Biochim. Biophys. Acta 1784 (2008) 973–982.
[13] W. Wei, J.H. McCusker, R.W. Hyman, T. Jones, Y. Ning, Z. Cao, Z. Gu, D. Bruno, M. Miranda, M. Nguyen, J. Wilhelmy, C. Komp, R. Tamse, X. Wang, P. Jia, P. Luedi, P.J. Oefner, L. David, F.S. Dietrich, Y. Li, R.W. Davis, L.M. Steinmetz, Proc. Natl. Acad. Sci. USA 104 (2007) 12825–12830.
[14] T. Jones, N.A. Federspiel, H. Chibana, J. Dungan, S. Kalman, B.B. Magee, G. Newport, Y.R. Thorstenson, N. Agabian, P.T. Magee, R.W. Davis, S. Scherer, Proc. Natl. Acad. Sci. USA 101 (2004) 7329–7334.
[15] U. Hoja, C. Wellein, E. Greiner, E. Schweizer, Eur. J. Biochem. 254 (1998) 520–526.
[16] M. Hablacher, A.S. Ivessa, F. Paltauf, S.D. Kohlwein, J. Biol. Chem. 268 (1993) 10946–10952.
[17] U. Hoja, S. Marthol, J. Hofmann, S. Stegner, R. Schulz, S. Meier, E. Greiner, E. Schweizer, J. Biol. Chem. 279 (2004) 21779–21786.
[18] R. Stucka, S. Dequin, J. Salmon, C. Gancedo, Mol. Gen. Genet. 229 (1991) 307–315.
[19] M.E. Walker, D.L. Val, M. Rhode, R.J. Devenish, J.C. Wallace, Biochem. Biophys. Res. Comm. 176 (1991) 1210–1217.

[20] N.K. Brewster, D.L. Val, M.E. Walker, J.C. Wallace, Arch. Biochem. Biophys. 311 (1994) 62–71.
[21] H.S. Kim, U. Hoja, J. Stolz, G. Sauer, E. Schweizer, J. Biol. Chem. 279 (2004) 42445–42452.
[22] B. Ng, S.W. Polyak, L.M. Bailey, D. Bird, J.C. Wallace, G.W. Booker, Anal. Biochem. 376 (1) (2008) 131–136.
[23] P.M. Glee, P.J. Russell, J.A. Welsch, J.C. Pratt, J.E. Cutler, Anal. Biochem. 164 (1987) 207–213.
[24] C. Cagnon, V. Valverde, J.M. Masson, Protein Eng. 4 (1991) 843–847.
[25] T. Vernet, D. Dignard, D.Y. Thomas, Gene 52 (1987) 225–233.
[26] R.J. Rothstein, Methods Enzymol. 101 (1983) 202–211.
[27] S. Jitrapakdee, K.H. Surinya, A. Adina-Zada, S.W. Polyak, C. Stojkoski, R. Smyth, G.W. Booker, W.W. Cleland, P.V. Attwood, J.C. Wallace, Int. J. Biochem. Cell Biol. 39 (2007) 2120–2134.
[28] K. Kowalski, A.L. Merkel, G.W. Booker, J. Biomolec. NMR 29 (2004) 533–534.
[29] S.W. Polyak, A. Chapman-Smith, P. Brautigan, J.C. Wallace, J. Biol. Chem. 274 (1999) 32847–32854.
[30] S.W. Polyak, A. Chapman-Smith, T.D. Mulhern, J.E. Cronan Jr., J.C. Wallace, J. Biol. Chem. 276 (2001) 3037–3045.
[31] D.F. Barker, A.M. Campbell, J. Mol. Biol. 146 (1981) 469–492.
[32] A. Chapman-Smith, T.D. Mulhern, F. Whelan, J.E. Cronan Jr., J.C. Wallace, Protein Sci. 10 (2001) 2608–2617.
[33] D.L. Val, A. Chapman-Smith, M.E. Walker, J.E. Cronan, J.C. Wallace, Biochem. J. 312 (1995) 817–825.
[34] A. Chapman-Smith, T.W. Morris, J.C. Wallace, J.E. Cronan, J. Biol. Chem. 274 (1999) 1449–1457.
[35] J.E. Cronan, Jnr, J. Biol. Chem. 265 (1990) 10327–10333.
[36] A. Chapman-Smith, D.L. Turner, J.E. Cronan, T.W. Morris, J.C. Wallace, Biochem. J. 302 (1994) 881–887.
[37] I. Chen, Y.A. Choi, A.Y. Ting, J. Am. Chem. Soc. 129 (2007) 6619–6625.
[38] E. Campeau, R.A. Gravel, J. Biol. Chem. 276 (2001) 12310–12316.
[39] A. Leon-Del-Rio, R.A. Gravel, J. Biol. Chem. 269 (1994) 22964–22968.
[40] L.M. Bailey, R.A. Ivanov, S. Jitrapakdee, C.J. Wilson, J.C. Wallace, S.W. Polyak, Hum. Mutat. 29 (2008) E47–E57.

Chapter 6

The characterisation of the domain structure of yeast biotin ligase and its complexes by small-angle X-ray scattering and molecular modelling.

This chapter describes the cloning, expression and purification of *S. cerevisiae* BPL (ScBPL). This allowed for the purification of large amounts of the enzyme leading to crystallisation trials as well as small angle X-ray scattering (SAXS) analysis of the protein. Data of ScBPL in solution with and without biotin, ATP and magnesium chloride was collected by SAXS and a model of the enzyme's structure has been generated. This is the first structure of a eukaryotic BPL to be determined. To further this information, random mutagenesis was performed on the N-terminal domain and key residues that affected enzymatic function were identified. This study has been formatted and submitted to JBC on the 1st May 2009.

Chapter 6:

The amino terminal domain of a eukaryotic biotin ligase adopts an amidotransferase fold and plays an active role in catalysis

Author	Contribution	Signature
Nicole R. Pendini	Expressed and purified ScBPL in bacteria and yeast. Performed mutagenesis, activity assays and analysis. Assisted in manuscript preparation.	
Nathan P Cowieson	Conducted SAXS experiments, data collection, analysis and figures and assisted in manuscript preparation.	
Grant W. Booker	Provided intellectual discussion and assisted in manuscript preparation.	
Matthew C. Wilce	Provided key intellectual discussion and assisted in proofreading.	
John C. Wallace	Provided intellectual discussion and assisted in manuscript preparation.	
Steven W. Polyak	Cloned and constructed expression plasmid. Provided discussion and key enzymology, assisted in manuscript preparation and editing of drafts.	

CHARACTERISATION OF THE DOMAIN STRUCTURE OF YEAST BIOTIN LIGASE AND ITS COMPLEXES BY SMALL-ANGLE X-RAY SCATTERING AND MOLECULAR MODELLING.

Nicole R. Pendini[1,2+], Nathan Cowieson[2,3+], John C. Wallace[1], Grant W. Booker[1], Matthew C. Wilce[2*] & Steven W. Polyak[1*].

[1]School of Molecular and Biomedical Science, University of Adelaide, North Tce, Adelaide, South Australia, Australia 5005

[2]Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria, Australia.

[3]Center for Synchrotron Science, Monash University, Clayton, Victoria, Australia.

+ Joint first authorship.

Running Title: *Saccharomyces cerevisiae* biotin protein ligase

KEYWORDS: yeast, *Saccharomyces cerevisiae*, biotin ligase, structure, SAXS.

Address correspondence* to: Matthew C. J. Wilce, Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria, Australia. Phone +61-3-9905-1086 Fax: +61-3-9905-3723

Email: matthew.wilce@med.monash.edu.au.

Address: Steven W. Polyak, School of Molecular and Biomedical Science, University of Adelaide, North Tce, Adelaide, South Australia, Australia 5005. Phone +61 8 8303 5289 Fax +61 8 8303 4362

Email: steven.polyak@adelaide.edu.au.

Eukaryotic biotin protein ligase (BPL) plays an essential role in gluconeogenesis and lipogenesis. There is a growing database of structural and biochemical data available for microbial BPLs, however the same is not true of eukaryote BPLs. The paucity of structural information is primarily due to difficulties in producing the large quantities of enzyme required for these studies. Here we present improved methods for the production of BPL from the yeast *Saccharomyces cerevisiae* (ScBPL). This has facilitated the first structural characterisation of a large BPL from a eukaryote using small-angle X-ray scattering (SAXS) and circular dichroism (CD). We propose a molecular model for ScBPL based upon homology identified with glutamine amidotransferase at the N-terminus and microbial BPLs at the C-terminus. SAXS measurements for the unliganded and liganded forms of ScBPL show a difference in gyroscopic radius, indicating the distance between the centre of mass of the two domains changes when ligand is introduced. CD measurements indicate a clear difference in the amount of α-helix and β-sheet between these different species of ScBPL. We propose this change is due to altering the conformation of the flexible linker to bring these domains in to close proximity for catalytic function. A novel role for the N-terminal domain in catalysis is proposed.

Biotin is a water-soluble vitamin that is activated and covalently attached to biotin-requiring enzymes by biotin protein ligase (BPL). BPL achieves biotinylation in a 2-step reaction in an ATP-dependent manner (Reaction 1 and 2). The structures of three BPLs have been solved and published, two from bacteria including *E. coli* (1) and *Staphylococcus aureus* (2) and one from the archae *Pyrococcus horikoshii* (3). These enzymes are comprised of a core catalytic module that appears to be conserved across BPLs from all species. This conservation is such that BPLs are interchangeable between organisms (4,5) allowing for cloning and analysis of BPLs by functional

complementation utilising a temperature-sensitive *E. coli* strain as host (6).

Biotinylation by the BPLs is a tightly controlled process whereby a specific lysine residue is recognised and modified by the covalent attachment of biotin to activate and act as a cofactor in carboxylation, decarboxylation and transcarboxylation reactions (4). The structure of the complex of *P. horikoshii* BPL and a substrate protein biotin carboxyl carrier protein suggests that specificity in the archaeal BPLs is achieved via an extended interaction interface on the conserved catalytic module (7). Biotin has also been found to control gene expression in both bacteria and eukaryotes (8,9). The presence of biotin in many bacterial BPLs leads to homodimerisation. This homodimer can then bind to the biotin biosynthetic operon via the N-terminal domains to cause repression of biotin synthesis (10). As eukaryotes do not synthesise biotin, the same control cannot be explained by the N-terminal of eukaryotes which show no sequence similarity to this or any other type of DNA binding motif.

Eukaryotic BPLs are approximately twice the size common to bacterial BPLs. They contain a conserved catalytic module at their C-terminal end while the structure and function of the N-terminal portion of eukaryotic BPLs remains unknown. N-terminal truncation studies of both *S. cerevisiae* and *H. sapiens* BPL show that N-terminal domains are required for catalysis in a mechanism that is not yet understood (11,12). Similarly, point mutations in the N-terminal domain of mammalian BPLs (known as holocarboxylase synthese (HCS)) lead to the disease, multiple carboxylase deficiency. This disease is caused by a failure of HCS to biotinylate its substrate proteins (12,13) implying the N-terminal is involved and required for the catalytic activity of mammalian BPL. This differs from its bacterial counterpart, where the much smaller N-terminal domain is only thought to play a significant role in the suppression of biotin synthesis.

Structural and biochemical studies of eukaryotic BPLs have been hampered by the difficulty of producing milligram amounts of full-length protein. Previous attempts to purify ScBPL for analytical and structural studies have resulted in great losses of full-length enzyme when expression was conducted in *E. coli* (11). In the present study a novel method to produce milligram quantities of stable, active ScBPL has been identified and used in combination with molecular modeling, small-angle x-ray scattering, circular dichroism and biochemical assays to probe the structure and function of the N-terminal domain of *S. cerevisiae* BPL (ScBPL).

Experimental Procedures

Molecular Modeling- The amino acid sequence of *Saccharomyces cerevisiae* BPL was used with the Phyre fold recognition server (14) for structure prediction analysis. Models generated by Phyre via a threading procedure to the best matches for the N and C terminal domains (namely PDBID: 1Q7R for the N-terminus and PDBID: 1HXD for the C-terminus) were used for further analysis..

DNA Manipulations- The expression vector pET16b (Novagen) was modified for the production of proteins containing an N-terminal c-myc epitope tag. pET-16b was restricted with *Nco*I and BamHI endonucleases and ligated with oligonucleotides B119 (5'-CATGGGTGAACAGAAACTGATCAGCGAAGAAGACCTCGAGTATACTGCAGG-3') and B120 (5'-GATCCCTGCAGTATACTCGAGGTCTTCTTCGCTGATCAGTTTCTGTTCACC-3'), thereby yielding pET-cmyc. The coding region for ScBPL was subsequently introduced into this vector by a three fragment ligation requiring: 1. *Xho*I / *Bam*HI treated pET-cmyc, 2. a 1.44 kB *Eco*RI / *Bam*HI fragment excised from pET(ScBPL-H_6) (11) and 3. a 800 bp fragment, obtained by PCR with

oligonucleotides B121 (5'-CTCGAGAACCTGTATCAGGGTATGAATGTATTAGTCTATAATGGC-3') and B109 (5'-CCAAAAGGCACATGCACAATCATATTCAACCCAGTTTTGGTAAGAAC-3').

pET(ScBPL-H_6) was treated with *Xho*I and *Eco*RI. This yielded pET(cmyc-TEV-ScBPL-H_6). A 2 kb fragment containing the coding region cmyc-TEV-ScBPL-H_6 was subsequently excised from this construct with *Nco*I and *Bam*HI restriction endonucleases and ligated into similarly treated pC104 vector (15), a derivative of pKK223-3 (Pharmacia). This produced the construct pK(cmyc-TEV-ScBPL-H_6).

Screening for N-terminal mutations- A library of random point mutations in the N-terminal domain of ScBPL-H_6 was obtained by error prone PCR using the procedure described in (15). Oligonucleotides B130/21 (5`-CGAGAACCTGTATTTCCAGGG-3`) and B112/24 (5`-CATCGTAAGAGGCTCTATAGCCTC-3`) were employed with template pK(cmyc-TEV-ScBPL-H_6) together with Taq DNA polymerase. PCR products were treated with *Xho*I and *Sac*I restriction endonucleases and ligated into similarly treated pK(cmyc-TEV-ScBPL-H_6). Plasmids were subsequently transformed into *E. coli* BM4062 (temperature sensitive *birA85* mutant strain) and grown at 30^{o}C for 16 hours. 1000 colonies were individually restreaked and grown at both 30^{o}C and 42^{o}C. Colonies that did not grow at 42^{o}C were the selected for further analysis. The integrity of expression plasmids was confirmed by colony PCR analysis using oligonucleotides B130/21 5'-TGCTGAAGGATAACACCGTGCCAC-3` and B76/21 5'-GAGTGCCTGCAGCTTTTCTGC-3`.

Expression of ScBPL-H_6 was confirmed in whole cell lysates by an ELISA using Ni-NTA HisSorb plates (Qiagen) and probing with a mouse anti-c-myc primary antibody then a sheep anti-mouse secondary antibody conjugated to HRP. Detection was achieved by addition of 0.01 mg mL^{-1} tetramethylbenzidine (TMB) in TMB buffer (10 mM sodium acetate, 10 mM EDTA) and allowed to develop for 15 minutes. The reaction was stopped with the addition of 100 μL 1 M sulphuric acid and measured at 450 nm. Plasmids from those colonies that expressed ScBPL-H_6 were subjected to DNA sequencing to identify the missense mutations.

Cloning and transformation of pVT(ScBPL-H_6) in Saccharomyces cerevisiae w303- pET(ScBPL-H6) (11)was treated with *Xba*I and *Xho*I endonucleases and the 2 kb fragment ligated into a similarly treated pVT100 (16). pVT(ScBPL-H_6) was transformed into competent *Saccharomyces cerevisiae* w303 cells as described in (17).

Expression and treatment of pVT(ScBPL-H6) - ScBPL expression was performed as described in (18) with the following modifications: ScBPL was resolved from residual contaminating proteins by anion exchange chromatography. A 15 mL Q-Sepharose column (Pharmacia) was equilibrated in ScBPL storage buffer (50 mM Tris-HCL pH 8.0, 0.5 mM EDTA pH 8.0, 5 % glycerol,1 mM DTT) using an AKTA FPLC system (Amersham Biosciences). IMAC purified and pooled proteins were loaded at 1 ml min^{-1} and fractionated with 60 mL gradient from 0 -200 mM NaCl in ScBPL storage buffer followed by 20 mL gradient from 200 mM -1 M NaCl in ScBPL storage buffer run at 1.5 ml min^{-1}. 1.5 mL fractions were collected and analysed by SDS-PAGE. All proteins were stored at -80^{o}C unless specified.

Activity assay for S. cerevisiae-expressed ScBPL- Measurement of BPL activity was determined based upon methodology described previously (11). Here, the procedure was modified such that scintillation counting could be performed in a 96 well format using a Packard Topcount scintillation plate reader (Perkin Elmer). The reaction mixture contained 50 mM Tris pH 8.0, 3 mM ATP, 5.5 mM

$MgCl_2$, 4.5 μM biotin, 0.1 mg mL^{-1} BSA, 0.1 μM DTT, 0.5 μM ^{3}H-biotin and 10 μM apo-*S.cerevisiae* Pyruvate Carboxylase 104 (11) in a final volume of 19 μL. The reaction mixture was pre-incubated for 5 minutes at 37^{o}C before addition of 1 μL of BPL (at 0.11 μg / μL) . Aliquots of the reaction (4.5 μL) were removed at 5 minute intervals and spotted onto Whatman 3MM paper pre-treated with 10 mM biotin in 50 mM Tris pH 7.5 followed by 10 % trichloroacetic acid (TCA) and set in a 96-well tray. The filters were washed twice in ice cold 10% TCA followed by one wash in 100% ethanol then air-dried. Each well was covered with 50 μL "Optisafe" scintillant (Packard) and the acid-insoluble radioactivity measured. Each well was read for 1 minute. Data were analysed with PRISM (GraphPad Software).

SAXS analysis- ScBPL was concentrated to 10 mg mL^{-1} in ScBPL storage buffer containing 250 mM NaCl. The biotin sample contained 10 mM biotin in 50 mM Tris-HCl was added to ScBPL resulting in a final concentration of 7.5 mg mL^{-1} and the biotin + ATP sample contained 10 mM biotin in 50 mM Tris-HCL pH 8.0, 10 mM ATP and 5 mM $MgCl_2$ added to ScBPL resulting in a final concentration of 7.5 mg mL^{-1}. SAXS data were collected on a SAXSESS instrument with a CCD (Anton Paar). A slit width of 10 mm was used with 150 μl of sample in a flow cell (Anton Paar). An exposure time of 180 x 10 seconds was used and data collected in the Q range 0.0058 to 0.15 $Å^{-1}$ for ScBPL, ScBPL + biotin and ScBPL + biotin + ATP. Equivalent aggregation states between the three samples was ensured by comparing the total forward scatter I(0) normalized for protein concentration and molecular weight. P(R) functions were calculated using the GNOM program (19). Rigid body fitting to the ATP and biotin liganded sample was done using the BUNCH algorithm (19). Molecular models of the N- and C-terminal domains were used to prepare the starting model. Fitting was repeated five times using different random seed start points. Analysis of the molecular motions of the molecules was done using the programs RANCH to generate 10,000 starting models with random linkers and GAJOE to optimize 50 ensembles with 50 curves per ensemble, 2000 generations and 50 cycles of optimization.

Circular Dichroism (CD) analysis- CD data was measured on a J-810 instrument (Jasco) between 190 and 260 nm. The data collection parameters were 0.1 mm path length, 1 second accumulation time, 2 nm bandwidth, 1 nm period. Spectra are the average of 5 accumulations. The apo, biotin-bound and holo-ScBPL samples were prepared as for SAXS data collection. Secondary structure percentages were calculated using the CDPro software package and are the average results of the SELCON3, CDSTTR and CONTINLL algorithms.

RESULTS

ScBPL contains an inactive amidopeptidase domain that is essential for activity. While there is no direct structural information available for a eukaryotic BPL, it is possible to model two discrete domains with a high degree of confidence using a threading approach. Using the primary sequence of ScBPL with the Phyre fold recognition server (14) gives a match between the C terminal 321 amino acids of ScBPL and the birA protein, the *E. coli* biotin repressor (PDB accession code 1HXD). This hit has an E value of $3.1e^{-29}$ (100 % confidence) with a sequence identity of 19 %. Similarly the N-terminal 208 amino acids yields four hits to class I glutamine amidotransferases, the best confidence value for these hits 0.0099 to a plasmodial glutaminase (PDB accession code 2ABW). The hit with the highest sequence identity was a glutamine amidotransferase from *Bacillus stearothermophilus* (PDB accession code 1Q7R) and the model based on this structure was used for subsequent analysis. This hit has an E value of 0.017 (95% confidence interval)

with a sequence identity of 20%. A sequence alignment between ScBPL and the birA and amidotransferase proteins shows good conservation of secondary structure as predicted based on sequence versus the two homologous structures from the biotin repressor and the glutamine amidotransferase. The exception is the 'winged-helix' motif of the birA, suggesting that this region is not conserved in ScBPL (figure 1a).

A number of residues have been identified as key to the formation of biotinyl-5`-AMP from biotin and adenosine triphosphate. These include the mainly basic residues R48, R51, D104, K111 and R233 in *Pyrococcus horikoshii* BPL. R118, R121, D176, K183 and R318 are the equivalent residues in *E. coli*. Two of these important residues are not conserved in ScBPL, namely N427 and L669 (R121 and R318 in *E. coli* BPL, figure 1a). In the structure of *Pyrococcus horikoshii* in complex with Biotin + ADP and Biotin + ATP (PDBID 2E1H and 2DTO respectively) show that these side-chains are positioned to interact with the negative charge of the phosphate of ADP/ATP. As ScBPL is missing these important conserved residues, it is plausible to suggest the N-terminal domain may provide the missing positively charged residues required for catalysis.

Amidotransferases catalyse the hydrolysis of glutamine to glutamate and ammonia via a Cys-His-Glu catalytic triad. In the case of ScBPL these catalytic residues are not conserved, and so this domain is unlikely to have amidotransferase catalytic activity. It may simply provide a binding site for a protein or small molecule. To probe the function of the putative amidotransferase domain of ScBPL random mutagenesis was performed. Mutants of cmyc-TEV-ScBPL-H_6 were screened by complementation in the *E. coli* strain BM4062 that contains a temperature-sensitive BirA mutant. Those isolates that were unable to be rescued by ScBPL at 42^{O}C were tested for full-length expression by binding of the C-terminal hexa-histidine tag to Ni-NTA HisSorb plates and by detection with c-myc antibodies binding the N-terminus. Three point mutants were identified in the N-terminal domain of ScBPL that reduced or abolished activity. These were L28P, L74S and G83D. Based upon the observed homology with glutamine amidotransferase, two of these mutations, L74S and G83D are within the N-terminal domain structure of this domain of ScBPL and are likely to disrupt the overall structure of ScBPL. However L28P is a surface exposed residue that lies in a loop region adjacent to the residues corresponding to the active site of the amidotransferase (figure 1b).

When the sequence of BPL from *S. cerevisiae* (690 residues) is compared to *H. sapiens* (726 residues) it is found that there is 20% sequence identity. Many naturally occurring mutations have been found in *H. sapiens* BPL (aka holocarboxylase synthetase, HCS) that lead to the disease multiple carboxylase synthetase deficiency (MCD) (13). These residues are found to be similar or identical to residues in ScBPL (figure 1a). Through the models generated for ScBPL, the mutations that perturb HCS activity have been mapped on the structure (figure 1b). Mutations at the N-terminal map to the opposite side of the amidotransferase active site and mutations in the C-terminal domain map around the known ligand binding site.

Over-expression of ScBPL in yeast yields large amounts of catalytically active protein. To facilitate structural analysis of ScBPL, an expression and purification strategy was developed. Expression of active and full length ScBPL in *E.coli* is hampered by proteolytic degradation and insolubility. Several methods were attempted to purify full length enzyme including mutagenesis of the proteolytic cleavage sites, K248T, expression of N-terminally (c-myc) and C-terminally (six histidine) tagged ScBPL to allow selective purification of full-length protein via these affinity tags (ie utilising the pK(cmyc-TEV-ScBPL-H_6) construct) and denaturing lysis condition with attempted refolding. All failed

to give substantial quantities of full length protein that remained active. However, over-expression of ScBPL in *S. cerevisiae* and purification by a combination of metal affinity and ion exchange chromatography yielded 9.6 mg of ScBPL-H_6 with purity greater than 95% from a 300 mL fermentation. Mass spectrometry on the purified enzyme indicates loss of 17 amino acids from the N-terminus $M^1NVLVYNGPGTTPGSVK^{17}$, however, the specific activity of this truncated protein was shown to be 36.7 nmol min^{-1} mg^{-1}, comparable to that obtained in (11) of 25 nmol min^{-1} mg^{-1} and significantly greater than the *E. coli* produced ScBPL-K248T protein which only had 5 nmol min^{-1} mg^{-1} activity. Catalytic activity of the purified ScBPL remained high 7 days post-purification indicating the stability of the protein produced by this strategy.

The biotin ligase and amidotransferase domains of ScBPL exist in both open and closed conformations and the equilibrium is altered by biotin and ATP. Small angle x-ray scattering was used to study the relationship between the two domains of ScBPL. The three samples of ScBPL used in this study were found to have constant values consistent with monomeric protein and do not show evidence of changes in aggregation state between samples (table 1). Indirect Fourier transformation of the data (20) results in a distribution function P(R) that represents the distribution of interatomic vectors within the particles (figure 2a). The P(R) plots for the three samples are significantly different indicting structural rearrangement between the apoform of ScBPL, the biotin and biotin-ATP liganded forms. The three curves contain strong peaks with maxima at ~ 40 Å that approximate Gaussian curves. The apo ScBPL is slightly skewed towards larger distances, consistent with a prolate sphere with an average radius of approximately 40 Å. Interestingly, as models of the biotin protein ligase and amidopeptidase domains approximate spheres of radius 20 Å, this peak is consistent with a juxtaposition of these two domains. The three curves also contain peaks of less intensity with maxima at 120 Å suggesting a smaller population of much longer vectors possibly consistent with an extended structure. The latter peak is more intense in the ScBPL + biotin sample than in ScBPL alone or ScBPL + biotin +ATP. In addition the unliganded ScBPL sample shows a distinct shoulder on the upper edge of the 40 Å maximum suggesting an intermediate population between the compact and extended structures.

To investigate the possibility that the two domains are in dynamic equilibrium between open and closed conformations, the SAXS data was analysed using the ensemble optimization method (21). This form of analysis generates a large pool of starting structures by linking the domains of known structure with linkers of the correct length and random orientations. A genetic algorithm is then employed to create ensembles of structures from the random pool that together are consistent with the measured SAXS data.

After analysis by the ensemble optimization method all three samples show discrete, normally distributed populations of compact and extended conformations. To visualize the various populations of conformations in the three samples, the optimized ensembles of protein structures were compared. The biotin ligase domain from each of the 50 conformations in each ensemble were aligned and a sphere was used to mark the center of mass of each of the amidopeptidase domains surrounding the ligase domain (figure 2b). Fifty models chosen at random from the unoptimised pool show a broad distribution of the amidotransferase domain around the ligase domain (yellow spheres, figure 2b (I). In contrast, the amidotransferase domain of the unliganded ScBPL clusters loosely into two populations one close to the biotin ligase domain, the other in extended conformations (cyan spheres, figure 2b (II). ScBPL with biotin has a similar distribution with many fewer molecules adopting the open

conformation (green spheres, figure 2b (III). In the closed conformation forms of ScBPL and ScBPL with biotin, the closed conformation amidotransferase domains seem to cluster at random either on the active site face of the biotin ligase or on the face opposite the active site. Interestingly, in the ATP containing sample the amidotransferase domains in the closed conformation strongly favours the face opposite to the active site and have few molecules in the extended conformation (magenta spheres, figure 2b (IV).

The linker between the two domains adopts a partially globular conformation and may drive the dynamics between the open and closed forms. Analysis of the relationship between the two domains of ScBPL show that in the presence of biotin + ATP, the amidotransferase domain clusters around the face of the biotin ligase domain that is distal to the active site. This tight clustering suggests that this liganded state may be largely in a single conformation and that models produced by simple rigid body fitting of the domains to the SAXS data are valid. Rigid body analysis refines both the relative positions of the two domains and also the path of an alpha carbon trace of the linker region against the SAXS data to produce a single model. Convergence of independent solutions of the modeling and good fit to the experimental data give confidence in the resulting models.

Five independent models of full length ScBPL were generated by the program BUNCH (19) against the ATP and biotin liganded sample. Chi values describing the fit of the five models to the data were in the range of 1.21 to 1.28 and a representative fit is shown (figure 3a). The amidotransferase domain of four of the five structures cluster around the face of the biotin ligase domain that is distal to the BPL active site while one clusters on the active site face (figure 3b). This is in good agreement with the model suggested by the ensemble optimization analysis. Interestingly, in all five models, the linker between the two domains refined in a globular arrangement with the bulk of density close to the amidotransferase domain and a small globular patch of density close to the ligase domain (figure 3c). The relatively globular conformation of the linker region is in agreement with the secondary structure predicted in this region (figure 1a).

To examine conformational change upon ligand binding, circular dichrosim was used to study the secondary structure of ScBPL in apo form and with biotin or biotin + ATP. From secondary structure prediction ScBPL should contain approximately 26.7 % helix, 24.1 % sheet and 49.3 % random structure. Secondary structure predictions based upon the CD data are summarised in table 2. Results from CD spectra (figure 3d) show there is indeed a distinct difference between apo, biotin bound and biotin + ATP ScBPL.

DISCUSSION

We present the results of molecular modeling analysis that suggests that the eukaryotic biotin protein ligase of yeast contains N- and C-terminal domains with homology to a class I amidotransferase and to the *E. coli* biotin-inducible repressor respectively. The lack of conservation of the catalytic triad of the putative amidotransferase domain suggests that it is catalytically inactive. However, results of the complementation assay conducted with several point mutations in the N-terminal region as well as truncation experiments conducted by Polyak and co-workers (11) demonstrate that this domain is essential to the function of ScBPL.

A novel expression and purification strategy for ScBPL is presented and has allowed for the first structural analysis of a eukaryotic BPL. Analysis by small-angle X-ray scattering show that the two domains of unliganded ScBPL exist in an equilibrium between an extended conformation and a closed conformation in which the amidotransferase domain is situated proximally to the ligase domain. Interestingly, the

presence of biotin moves the equilibrium in favour of the closed conformation. However, in this case, the amidotransferase domain does not cluster evenly around distal faces of the ligase domain implying that there are two sites of interaction between the biotin ligase and amidotransferase domain that may be mediated by a partially folded arrangement of the linker region. Secondary structure prediction and rigid body fitting of the linker region to SAXS data is consistent with a structured and globular linker region.

Addition of ATP to the biotin liganded ScBPL results in further compaction of the protein (a reduction in radius of gyration from 42 to 36 Å) and clustering of the amidotransferase domain to the distal side of the biotin ligase domain. We speculate that binding of ATP may result in an interaction between the amidotransferase domain and the face of the biotin ligase domain distal to the active site. Such a model has a precedent in a structurally related amidotransferase domain containing protein. CTP synthase contains two domains, an N-terminal synthase domain responsible for the synthesis of CTP and a C-terminal amidotransferase domain. In this example, binding of GTP to the CTP synthase domain causes a conformational change that is thought to orient the amidotransferase domain such that it is in a favourable position to couple glutamine deamidation with CTP synthesis (22). Thus the binding of a ligand to the synthase domain of CTP synthase causes recruitment of the amidotransferase domain and allosteric activation of CTP synthesis. Thus, nucleotide-driven recruitment to neighboring catalytic domains may be a common feature of amidotransferase domains.

Based on the results of SAXS analysis we propose a model in which the linker region between the catalytically inactive amidotransferase domain of yeast BPLs is in equilibrium between an unfolded and folded state. Addition of biotin to the enzyme moves this equilibrium towards the closed state, where the linker becomes more structured and the two domains are close together in space. Addition of biotin and ATP causes a conformational change in the ligase domain that allows the amidotransferase domain to interact with the face of the ligase domain distal to the active site, causing an allosteric activation and stabilisation of the biotin ligase domain in ScBPL. The linker is now more unstructured to allow the release of the reactive intermediate biotinyl-5'-AMP by the dissociation of the N- and C-termini and may be under strain due to the compressed nature of the domains (figure 4). As this enzyme's role is to form biotinyl-5'-AMP, it is tempting to propose that, when biotin is added, the N-terminal moves close to the ATP binding site. The many lysine and arginine residues on the surface of the N-terminal of ScBPL may stabilise the negative charge of the β and γ phosphate groups of ATP, providing the basic charge that is not present in the active site of ScBPL but is observed to be essential in bacterial and archaeal homologues. Once the biotinyl-5'-AMP is formed, the overall negative charge is reduced as the P-P_i is expelled and the N-terminal relaxes to a third confirmation.

From these results it has been confirmed for the first time that ScBPL has two structured domains and there is evidence that the linker contains structured regions within it that are responsible for mediating the proximity between the N- and C-termini of ScBPL. Mutations in the N-terminal of ScBPL map around the active site of the amidotransferase active site suggesting this is the side which interacts with the catalytic domain of ScBPL. However, sequence homology to HCS suggests the opposite side to the amidotransferase is involved in catalytic function. ScBPL crystallisation trials are currently underway to ultimately resolve the precise position of the domains, to determine which interfaces of each domain interact and the differences in the linker region between unliganded and liganded forms of ScBPL.

REFERENCES

1. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., and Matthews, B. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9257-9261
2. Pendini, N. R., Polyak, S. W., Booker, G. W., Wallace, J. C., and Wilce, M. C. (2008) *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 520-523
3. Bagautdinov, B., Kuroishi, C., Sugahara, M., and Kunishima, N. (2005) *J Mol Biol* **353**, 322-333
4. Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C., and Wood, H. G. (1988) *J. Biol. Chem.* **263**, 6461-6464
5. Reed, K. E., and Cronan, J. E. J. (1991) *J. Biol. Chem.* **266**, 11425-11428
6. Uchida, K. M., and Otsuka, A. J. (1987) *Mol Gen Genet* **210**, 234-240
7. Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2008) *J Biol Chem*
8. Dakshinamurti, K., and Desjardins, P. R. (1969) *Biochim Biophys Acta* **176**, 221-229
9. Rodriguez-Melendez, R., Lewis, B., McMahon, R. J., and Zempleni, J. (2003) *J Nutr* **133**, 1259-1264
10. Beckett, D. (2007) *Annu Rev Genet* **41**, 443-464
11. Polyak, S. W., Chapman-Smith, A., Brautigan, P. J., and Wallace, J. C. (1999) *J Biol Chem* **274**, 32847-32854
12. Campeau, E., and Gravel, R. A. (2001) *J Biol Chem* **276**, 12310-12316
13. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C., Wallace, J. C., and Polyak, S. W. (2008) *Biochim Biophys Acta* **1784**, 973-982
14. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J., and Kelley, L. A. (2008) *Proteins* **70**, 611-625
15. Polyak, S. W., Chapman-Smith, A., Mulhern, T. D., Cronan, J. E., Jr., and Wallace, J. C. (2001) *J Biol Chem* **276**, 3037-3045
16. Vernet, T., Dignard, D., and Thomas, D. Y. (1987) *Gene* **52**, 225-233
17. Rothstein, R. J. (1983) *Methods Enzymol* **101**, 202-211
18. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C., Wallace, J. C., and Polyak, S. W. (2008) *Arch Biochem Biophys* **479**, 163-169
19. Petoukhov, M. V., and Svergun, D. I. (2005) *Biophys J* **89**, 1237-1250
20. Konig, S., Svergun, D., Koch, M. H., Hubner, G., and Schellenberger, A. (1992) *Biochemistry* **31**, 8726-8731
21. Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007) *J Am Chem Soc* **129**, 5656-5664
22. Willemoes, M., Molgaard, A., Johansson, E., and Martinussen, J. (2005) *FEBS J* **272**, 856-864

Equation 1 and 2

1) biotin + MgATP + BPL ⟷ biotinyl - 5'- AMP + PP_i

2) biotinyl - 5'- AMP + apo-BCCP ⟶ biotinylated (holo)-BCCP + AMP

Table 1

sample	concentration (mg/ml)	molecular weight (Kda)	I(0)	constant	Rg (Å)
ScBPL	10	76.4	0.606	0.00079319	41.2
ScBPL + biotin	7.5	76.4	0.436	0.00076091	42.1
ScBPL + biotin + ATP	7.5	76.4	0.392	0.00068412	36.1

Table 1. Statistics from the ScBPL SAXS experiments. The sample concentrations in mg/ml, molecular weight in kDa, and scattering amplitude I(0) and radius of gyration (Rg) in angstroms (Å) are shown. The constant is equal to I(0)/concentration/molecular weight.

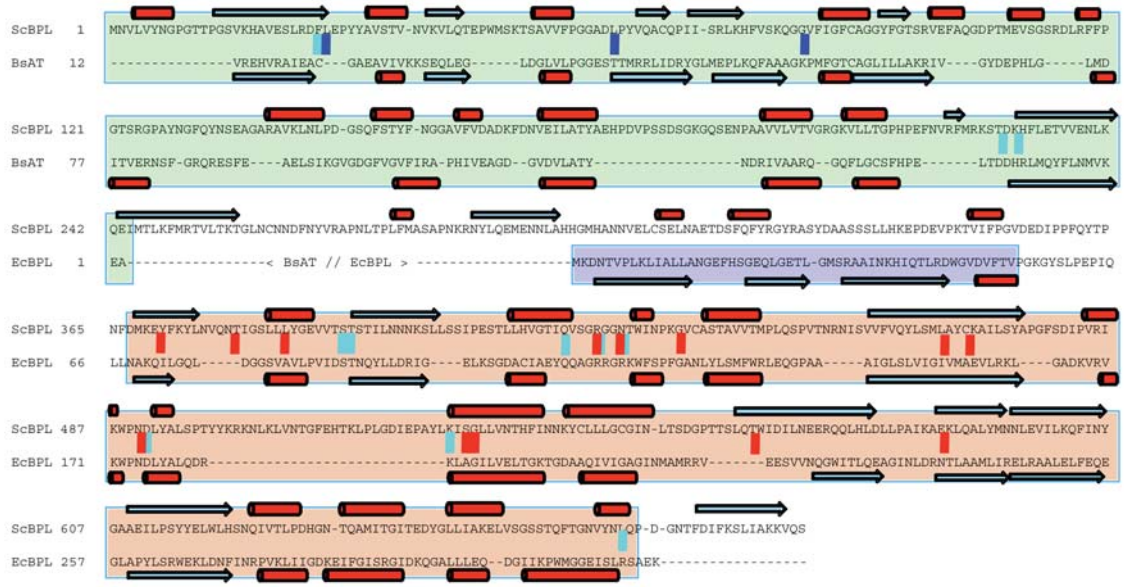
Table 2

sample	helix (%)	sheet (%)	random (%)
ScBPL	20.9	31.2	47.2
ScBPL + biotin	30.9	20.5	48.8
ScBpl + biotin + ATP	18.2	29.2	53.3

Table 2. ScBPL CD analysis
The percentages of α-helix, β-sheet and random structure calculated from CD data in protein solution samples of ScBPL, ScBPL+ Biotin and ScBPL + Biotin + ATP, as described in experimental procedures.

Figure 1

A


ScBPL 1 MNVLVYNGPGTTPGSVKHAVESLRDFLEPYYAVSTV-NVKVLQTEPWMSKTSAVVFPGGADLPYVQACQPII-SRLKHFVSKQGGVFIGFCAGGYFGTSRVEFAQGDPTMEVSGSRDLRFFP
BsAT 12 --------------VREHVRAIEAC---GAEAVIVKKSEQLEG------LDGLVLPGGESTTMRRLIDRYGLMEPLKQFAAAGKPMFGTCAGLILLAKRIV----GYDEPHLG-----LMD
ScBPL 121 GTSRGPAYNGFQYNSEAGARAVKLNLPD-GSQFSTYF-NGGAVFVDADKFDNVEILATYAEHPDVPSSDSGKGQSENPAAVVLVTVGRGKVLLTGPHPEFNVRFMRKSTDKHFLETVVENLK
BsAT 77 ITVERNSF-GRQRESFE----AELSIKGVGDGFVGVFIRA-PHIVEAGD--GVDVLATY----------NDRIVAARQ--GQFLGCSFHPE-------LTDDHRLMQYFLNMVK
ScBPL 242 QEIMTLKFMRTVLTKTGLNCNNDFNYVRAPNLTPLFMASAPNKRNYLQEMENNLAHHGMHANNVELCSELNAETDSFQFYRGYRASYDAASSSLLHKEPDEVPKTVIFPGVDEDIPPFQYTP
EcBPL 1 EA----------< BsAT // EcBPL > ----------MKDNTVPLKLIALLANGEPHSGEQLGETL-GMSRAAINKHIQTLRDWGVDVFTVPGKGYSLPEPIQ
ScBPL 365 NFDMKEYFKYLNVQNTIGSLLLYGEVVTSTSTILNNNKSLLSSIPESTLLHVGTIQVSGRGGNTWINPKGVCASTAVVTMPLQSPVTNRNISVVFVQYLSMLAYCKAILSYAPGFSDIPVRI
EcBPL 66 LLNAKQILGQL-----DGGSVAVLPVIDSTNQYLLDRIG----ELKSGDACIAEYQQAGRRGRKWFSPFGANLYLSMFWRLEQGPAA-----AIGLSLVIGIVMAEVLRKL----GADKVRV
ScBPL 487 KWPNDLYALSPTYYKRKNLKLVNTGFEHTKLPLGDIEPAYLKISGLLVNTHFINNKYCLLLGCGIN-LTSDGPTTSLQTWIDILNEERQQLHLDLLPAIKAEKLQALYMNNLEVILKQFINY
EcBPL 171 KWPNDLYALQDR----------KLAGILVELTGKTGDAAQIVIGAGINMAMRRV-------EESVVNQGWITLQEAGINLDRNTLAAMLIRELRAALELFEQE
ScBPL 607 GAAEILPSYYELWLHSNQIVTLPDHGN-TQAMITGITEDYGLLIAKELVSGSSTQFTGNVYNQP-D-GNTFDIFKSLIAKKVQS
EcBPL 257 GLAPYLSRWEKLDNFINRPVKLIIGDKEIFGISRGIDKQGALLLEQ--DGIIKPWMGGEISLRSAEK----------

B

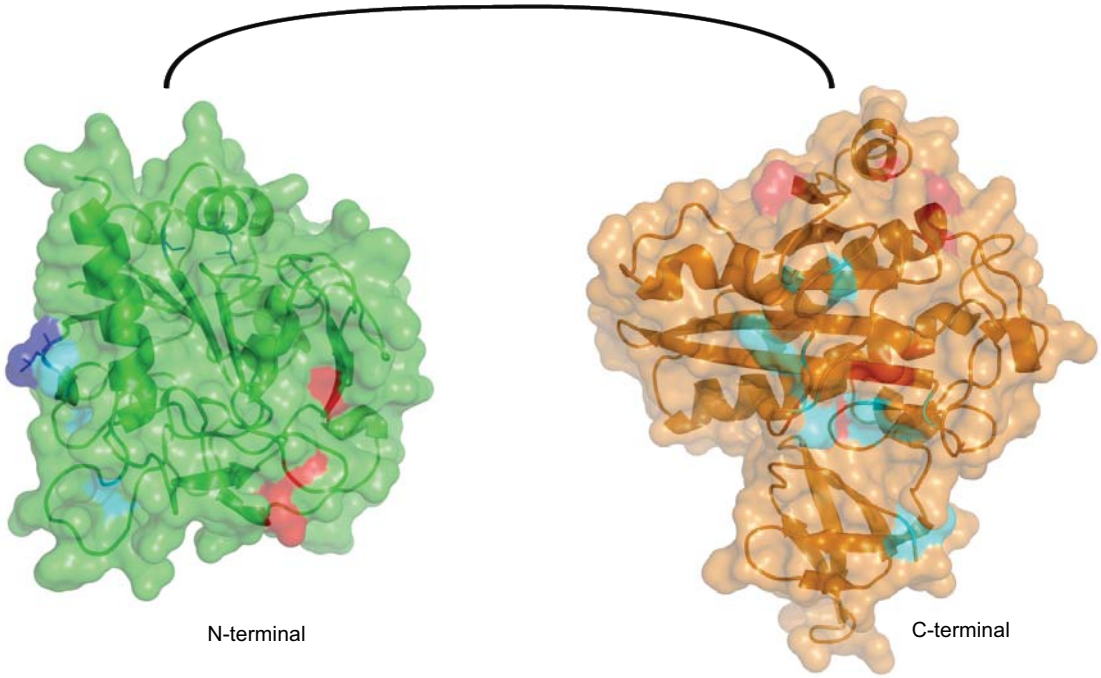

N-terminal
C-terminal

Figure 1 ***Sequence alignment and structural similarity of the N- and C-terminal domains of ScBPL to B.stearothermophilus amidotransferase and E. coli biotin ligase.***

Panel A shows a sequence alignment between ScBPL and an amidotransferase from *Bacillus stearothermophilus* (BsAT)(PDBID: 1Q7R) and the biotin inducible repressor protein from *Escherichia coli* (EcBPL)(PDB accession code 1HXD). The N-terminal domain of ScBPL is aligned with the amidotransferase and boxed in green while the C-terminal domain is aligned with the biotin ligase and boxed in orange. The 'winged-helix' DNA binding domain of the EcBPL is boxed in purple and is not present in the eukaryotic protein. The linker region of ScBPL shows no identity and therefore is not aligned with either protein. Secondary structure elements are represented by red cylinders (beta-sheet) and blue arrows(alpha-helix). The secondary structure elements for the two proteins of known structure are taken from their structure while the ScBPL secondary structure elements is predicted by the PHD algorithm (14). Residues corresponding to the active site cysteine, aspartate and histidine residues of the amidotransferase are marked by a cyan vertical bar. The activity reducing mutations identified by the complementation assay in the amidotransferase domain of ScBPL are show by vertical navy bars. Active site residues of the biotin ligase domain are marked by vertical cyan bars and mutations that lead to MCD in human BPL that are identical or similar in ScBPL are marked in red.

Panel B shows a surface and secondary structure representation of the ScBPL model. The biotin ligase domain is shown in orange and the amidotransferase domain in green. Active site residues of the amidotransferase at the N-terminal and the biotin ligase domain at the C-terminal are coloured cyan. The mutations derived from the complementation assay are coloured dark blue. A black line links the C-terminus of the amidotransferase domain to the ligase domain, MCD-causing mutations are shown in red. The two domains have been oriented to show the active site faces and do not represent their actual orientation.

Figure 2

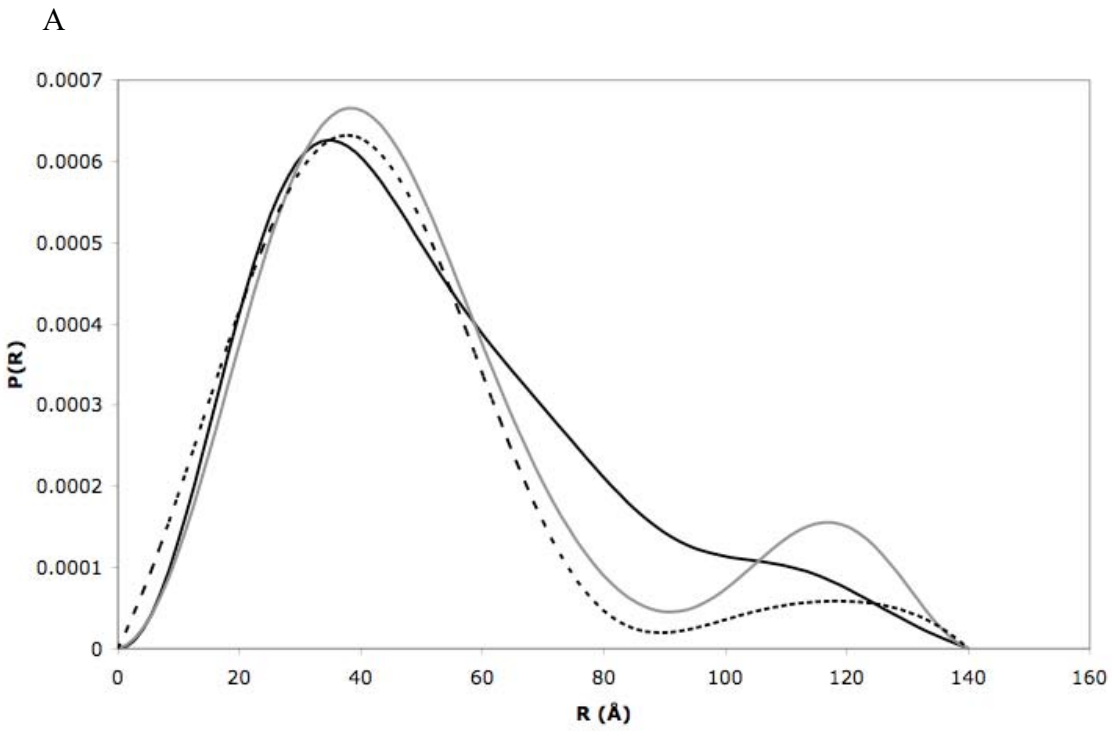

A
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0
P(R)
0
20
40
60
80
100
120
140
160
R (Å)
B
I
II
III
IV

Figure 2

A) Graphical representation of the distribution of interatomic vectors within ScBPL in various liganded states.

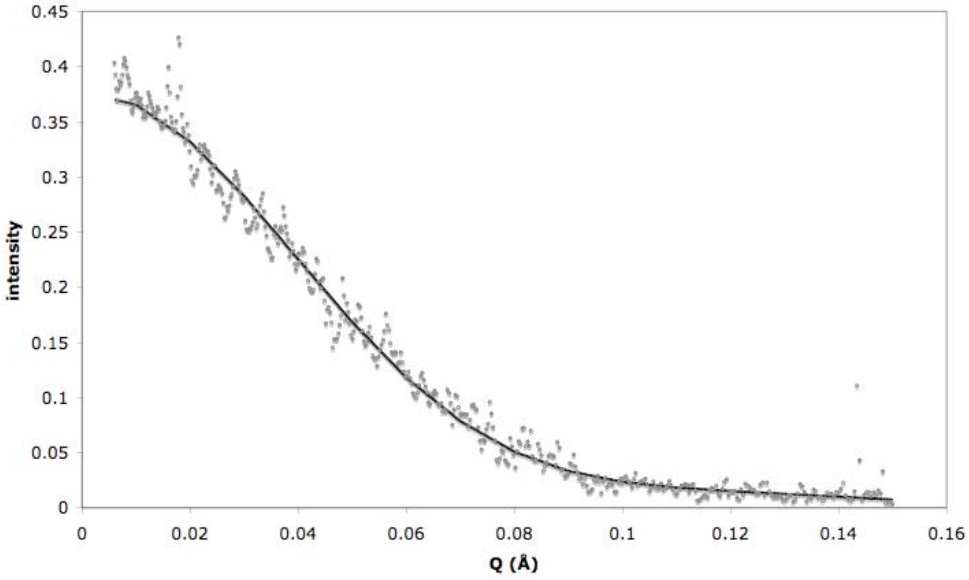
The probability density function P(R) was calculated from GNOM (19). Unliganded ScBPL is shown as a black solid line, ScBPL with biotin is shown as a grey line and ScBPL with biotin and ATP is shown by a dashed black line.

B) Domain conformations adopted by ScBPL in various liganded states.

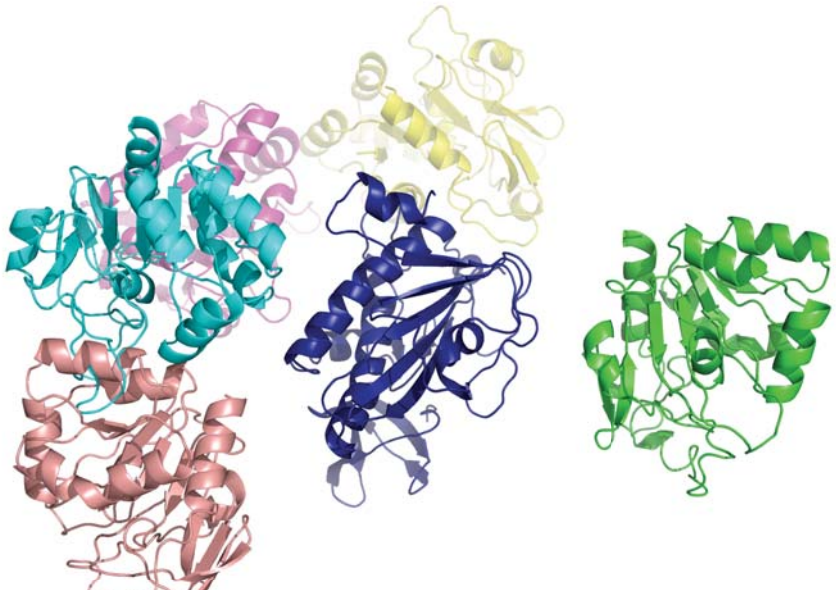
The modeled biotin ligase domain of ScBPL is in dark blue as a cartoon representation. The spheres surrounding the biotin ligase domain represent the center of mass of the amidotransferase domain of each of the 50 models in the ensemble of models that best fit the data. The yellow spheres are 50 models selected at random from the pool of models and represents the possible spread of conformations (I). Unliganded ScBPL is shown with cyan spheres (II), ScBPL with biotin with green spheres (III), and ScBPL with biotin and ATP with magenta spheres (IV). Black circles were overlaid on the figure as a visual aid to comparing the figures.

Figure 3

A


intensity
Q (Å)
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.02
0.04
0.06
0.08
0.12
0.14
0.16

B



C

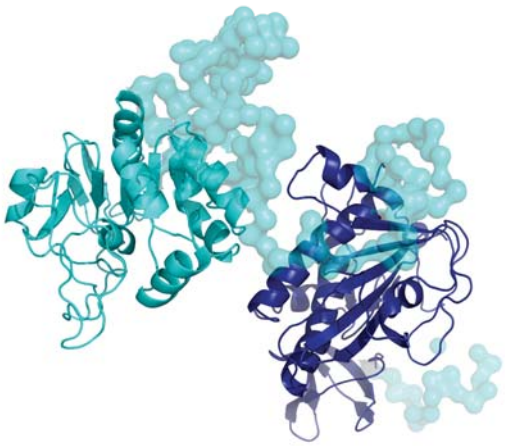


Figure 3

D

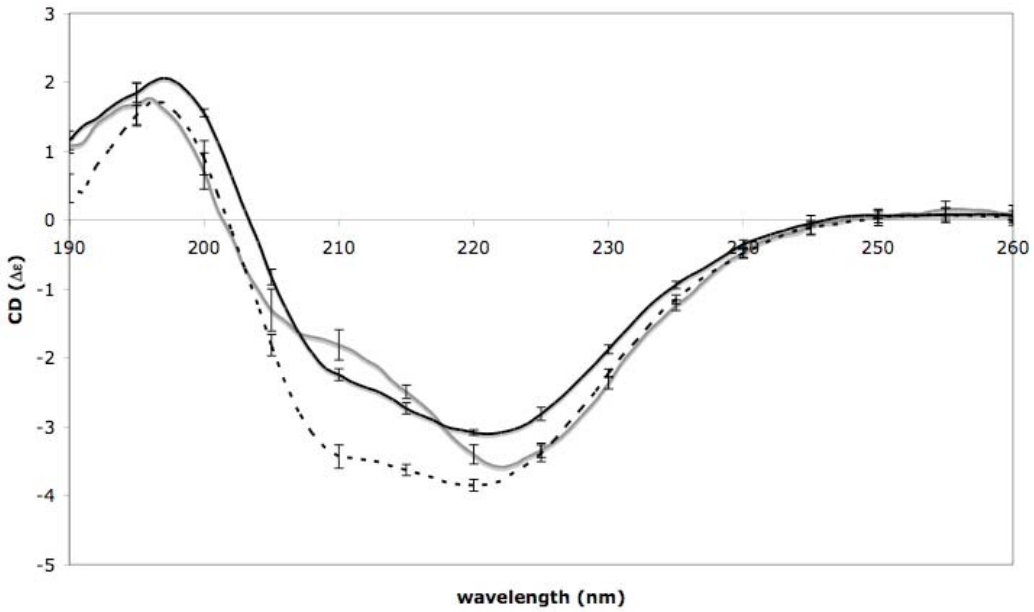


Figure 3 ***Rigid body modeling from the SAXS data.***
Panel A shows a representative fit to the data of one of the rigid body models.
Panel B shows 5 models of ScBPL generated by independent runs of the program BUNCH (19). The biotin ligase domain (shown in dark blue) of the 5 models were superimposed, and are shown with their active sites to the right hand side.
Panel C shows a representative model of ScBPL in the same orientation as panel B. The refined alpha carbon positions of the linker is shown in a surface representation.
Panel D *shows the CD spectra of ScBPL. Samples of ScBPL in unliganded and liganded were analysed by CD between wavelengths 190 and 260 nm. Apo ScBPL is represented by solid black line, ScBPL+ Biotin by black dashes and ScBPL+ Biotin+ ATP+*Mg^{+2} *by grey solid line. Error bars show standard error from 5 replicates.*

Figure 4

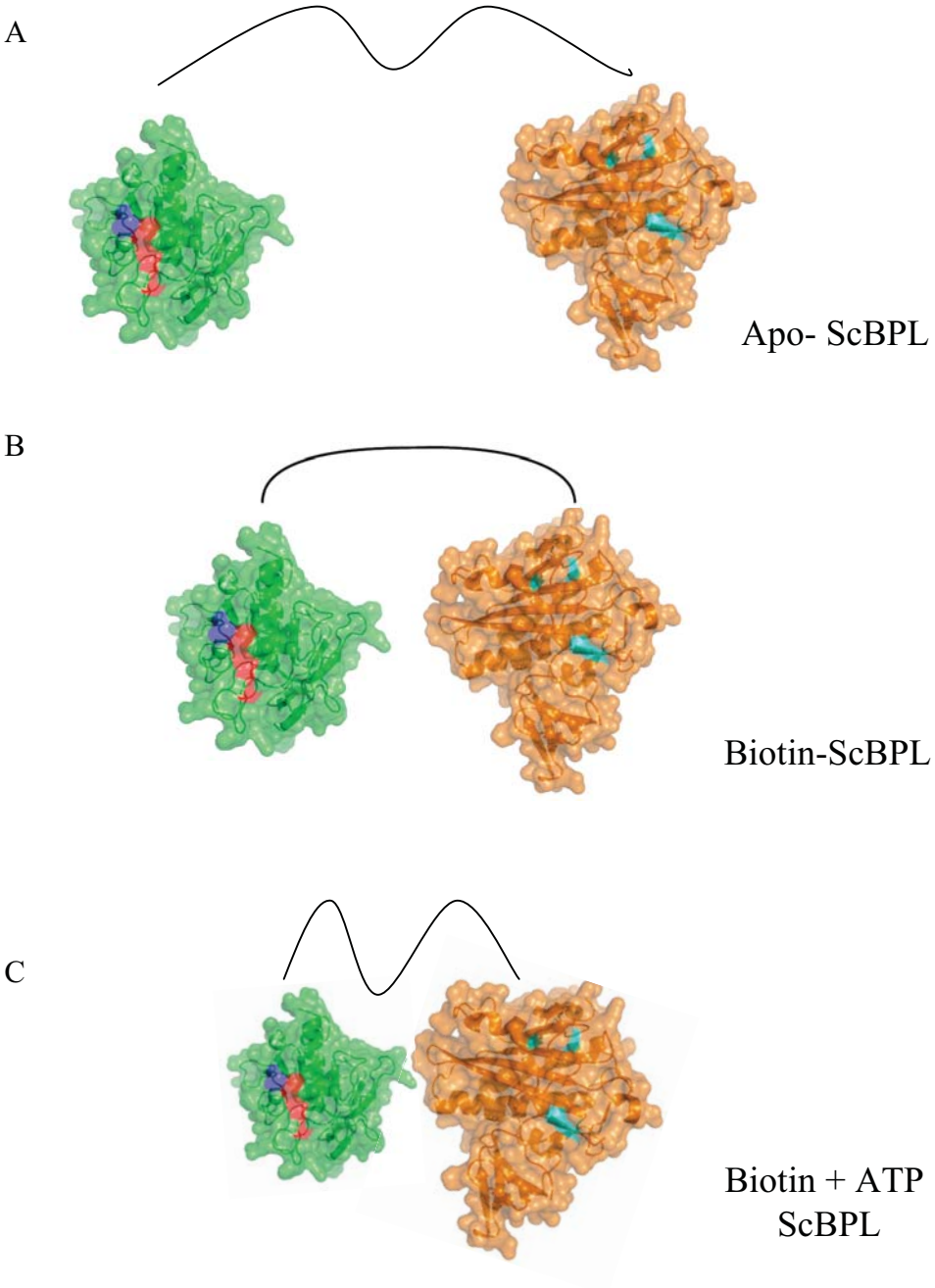


Figure 4 ***Models of ScBPL in unliganded and liganded conformations.*** Based upon the SAXS and CD data of ScBPL in the apo form, Biotin bound and Holo forms, 3 models have been generated to explain the differences between these states. A) Apo-ScBPL: the linker is more sheet like and extended, the domains are further apart in space, B) Biotin-ScBPL: the linker between the 2 domains is more helical and brings the 2 domains closer in space and C) Biotin + ATP- ScBPL: the 2 domains are compact and the linker is more sheet-like and unstructured .

Chapter 7

Discussion and future directions

In view of the essential role of BPL in all organisms, and the differences observed in the isozymes between the phyla, it has been proposed that BPL may be a target for anti-infective agents that specifically target a pathogen over the host. Two clinically important pathogens for which current therapeutics are becoming less effective due to resistance are *Staphylococcus aureus* and *Candida albicans*. Hence, specific focus has been applied to the BPLs from these species in this study in addition to host BPL from *Homo sapiens*.

SaBPL as a drug target for new anti-infective agents

The Gram-positive bacterium *S. aureus* was once susceptible to a wide range of antibiotics including penicillin, β-lactam, methicillin and its derivatives. However, this pathogen has evolved to resist all these therapies. Resistance has occurred due to indiscriminate prescription and the resulting over-use of antibiotics. These practices place selective pressure to favour bacterial isolates with adaptive features that enable avoidance of these drugs (65). Events that lead to resistance include transfer of mobile genetic elements from other bacteria, expression of proteins that bind the antibiotic and prevent binding to its desired target, thickening of cell walls such that drugs cannot penetrate, and spontaneous DNA mutation(s) that render the drug target resistant.

Antibiotics currently in use in the clinic target several major metabolic pathways including fatty acid, DNA, protein and folic acid synthesis. Due to the fundamental importance of BPL to the pathway providing the fatty acids required for membrane biogenesis, it has been proposed that BPL is a target for antibiotic development. Bacterial ACC has also been shown by Freiberg *et al.* to be a good novel target for inhibition (66).

Improvement of the *Staphylococcus aureus* BPL structure

In this study, the structures of various liganded states of SaBPL have been obtained at moderate resolution. There is an appreciation that improvement on this resolution will be sought in order to use these structures for structure-based drug design. One approach to obtain higher resolution structures

requires improving the protein for crystallography. For example, removal of the histidine tag used for purification could be pursued. To this end, a construct was prepared to place the His tag at the N-terminal of SaBPL with a Tobacco etch virus (TEV) cleavage site to facilitate the tag's removal. Expression and purification of the His_6-TEV-SaBPL was performed as per the SaBPL-His_6 construct (chapter 3) (67). Crystallisation trials were conducted with the His_6-TEV-SaBPL protein but no crystals were obtained. Removal of the His tag was successfully performed but crystallisation trials are still required.

Optimisation of the crystallisation conditions is another way of obtaining a high-resolution structure. Improved crystals could be obtained by growing crystals in the presence of various additives. Finally, improved data collection could be attempted by trying alternative cryo-protectants. Many trials were conducted around the initial crystallisation condition obtained from the Sigma kit, including altering concentrations of the precipitating reagents, pH, temperature and additive screens. From these methods one other condition was found to produce plate like crystals, and diffraction from these crystals was observed in one direction. However, diffraction deteriorated as the crystal was rotated and showed no improvement over data already obtained.

Further Development of Antimicrobials Targeting BPL

Future directions for the project include continuing the *in silico* screening trials that have already begun with SaBPL to identify potential leads as drug candidates. Some leads have already been discovered using *in silico* screening with the Schrodinger suite of programmes (http://www.schrodinger.com/) and the holo SaBPL structure. This can allow for direct docking to a selected target region of the protein (usually the active

site), of known compounds that are commercially available. The compounds may be part of a freely available library such as the National Cancer Institute suite of compounds (http://cactus.nci.nih.gov/ncidb2/download.html) for which if a hit was found, small amounts of the compound can be ordered for *in vitro* analysis against BPL. If BPL activity is inhibited by a compound, further optimisation and selectivity can be made to the compound to target SaBPL without affecting HCS. Another way of identifying lead compounds would be to search the ZINC database (http://zinc.docking.org/) for molecules that are of similar molecular weight and properties to biotin or ATP, known as core based matching, but could potentially have a higher affinity to the active site over the natural ligand.

Once such a hit has been identified by this basic docking methodology, a more advanced and intellectual method of drug design can be employed. A small molecule fragment library can be used and attached to various positions of the lead molecule based on "pockets" observed in the active or remote site of the target structure. In this study the subtle differences between various organisms in residues surrounding the active site of BPL may be used to develop high affinity selective inhibitors. New drugs that are currently in the pipeline to target *S. aureus* are mainly structural derivatives of those already on the market. They are easy to fast track through the drug discovery pipeline as they have a known mode of action and usually obtain FDA approval readily. However, problems with resistance can quickly arise thereby limiting the useful shelf-life. One solution to this problem is to target alternative pathways and design antibiotics unrelated to those that the bacteria have been exposed to. A major drawback with this approach is that it takes more research time and money.

However, our multidisciplinary team consisting of researchers from the University of Adelaide and Monash University has designed compounds to directly target the active site of SaBPL. We have shown so far that some of these compounds have *in vitro* activity against SaBPL and not HCS. Additionally, the compounds also show whole cell activity against *S. aureus*. Attempts to co-crystallise these compounds with the enzyme are progressing. Though there is much work yet still to be done for the development of this completely novel antibiotic, work presented in this thesis has greatly impacted on the strength of the project and has resulted in filing a PCT application describing the SaBPL structures and uses thereof for structure-based drug targeting.

Eukaryotic BPL structure & function – background

Prior to this study, structures had been published for *E. coli* and *P. horikoshii* BPLs (as summarised in chapter 2 (68)). However, there was very limited structural information known about eukaryotic BPLs. It has previously been shown that for HCS and for *S. cerevisiae* BPL, truncations to the N-terminal domain cause a decrease or drastically abolished enzymatic function (54,57). Furthermore L. Mayende *et al.* have performed limited proteolysis on HCS to map the structural domains, followed by yeast two-hybrid screens to demonstrate the interaction between the N- and C-terminal halves of the enzyme (69). Understanding how mutations to residues within the active site and catalytic core disturb function and decrease affinity for the ligand biotin is simplified by homology modelling and it has been shown that patients with these mutations respond relatively well to biotin supplementation (70). Structures of the well characterised bacterial and archaeal homologues of BPL

were shown to be very useful for modelling the structure of the catalytic region of HCS and positioning MCD mutations of the catalytic domain in a 3-dimensional fold, thereby providing a molecular explanation for impaired enzyme function. However, there is also a group of point mutations that cluster distal to the active site and impaired function by a mechanism that is not understood. Patients with these mutations respond poorly to biotin supplementation and have a poor long-term prognosis. One particular mutation, L216R, has been extensively characterised and shown to be essentially catalytically inactive. Our group has recently demonstrated that the mutant protein has a half-life that is half that of wild type HCS (71). There have been several studies reporting biochemical information on mutations of HCS that lead to MCD. However, no 3-dimensional structural data had been published detailing how and why these mutations lead to disease states.

Determination of Eukaryotic BPL structure

There is an appreciation that whilst models can be useful in understanding protein structure – function relationships, actual high-resolution structures are most preferred. The first, and often rate limiting, step in obtaining a structure is to obtain suitable quantities of purified protein for analysis. Several trials to purify full length HCS by recombinant expression in bacteria (BL21 and Rosetta cell lines) and in yeast (*S. cerevisiae* w303 and *Kluyveromyces lactis* cell lines) resulted in low yields of soluble protein and proteolytically degraded products. These expression systems require further optimisation in order to obtain the quantity of full-length enzyme required for crystallisation trials. One system for which expression trials were initiated was the *K. lactis* yeast secretion system. This approach uses an N-terminal signalling sequence to

target the protein out of the cell into the growth media. The C-terminal hexa-histidine tag could then be used to readily enrich HCS from the media. From small-scale trials, the media did contain product that bound NiNTA conjugate when analysed by Western-dot blot technique (72) indicating expression of protein containing the C-terminal His tag. Should up-scaling of protein expression not be successful in producing suitable amounts of full-length enzyme, a mammalian expression system or truncated constructs of HCS could be trialled. Shortened protein forms of the various predicted structural domains of HCS could allow for investigation by NMR. Techniques such as circular dichroism, electron microscopy and small angle X-ray scattering could also be used to estimate overall topology of the full-length enzyme.

The role of the Eukaryotic BPL N-terminal domain in disease

These efforts into understanding the structure of HCS will be useful to decipher the molecular mechanism of catalysis for the large eukaryotic BPLs. Already we observe residues from the N-terminal of the protein that, when mutated, impair enzyme activity. It is tempting to suggest that these residues could position themselves to interact with the active site or stabilise the structure or function of the catalytic domain. In addition to an understanding of why there is such a large difference between eukaryotic and prokaryotic BPL in terms of size, the mechanism of substrate recognition may also be answered. Eukaryotic BPLs in general have more substrates to recognise and biotinylate than do prokaryotic BPLs, and it is possible that the N-terminal extension may determine the order in which carboxylases are biotinylated in situations where biotin supply is limited. There may even be N-terminal post-translational modifications that alter the substrate recognition for eukaryotes. However, no

epitopes that match those known for other post translational modifications, including phosphorylation or glycosylation, were found when searches were conducted in ExPASy (http://au.expasy.org/tools/). There is however some evidence indicating that ScBPL itself may be biotinylated, as western blots probed with streptavidin-HRP detect full-length ScBPL (data not shown). Due to success with other BPL's at that time requiring all my efforts, this result was not further characterised. However, it would be interesting to digest the protein and perform mass spectrometry on the fragments to determine if in fact the enzyme is itself biotinylated. Why the enzyme itself would need to be biotinylated does not fit with biotin's conventional role as a co-factor for carboxylases. However, like its prokaryotic ancestor, eukaryotic BPL may also be a protein possessing a dual function which will require more analysis to truly decipher the role of the N-terminal domain.

It is imperative that we understand as much as we can about this enzyme so those patients who are poorly responsive to biotin therapy have alternative treatment remedies. Alternative approaches that could be investigated include designing biotin analogue(s) that have higher affinity for the mutant enzyme or higher catalytic activity when attached to the carboxylase. Knowledge about the structures could be used to design novel chemical chaperones to help stabilise the mutant proteins and protect it from turn over. Such a molecular “chaperone” has been described for phenylalanine hydroxylase (73),(74). High throughput screening assays developed here to discover inhibitors of BPL (75) could be employed to identify drugs that increase HCS activity.

Novel therapeutics through BPL targeting

Of the fungal pathogens, Candida spp. are the most predominant cause of invasive infection (76). In light of this fact, a similar pathway for drug design has been envisaged for *Candida albicans* using CaBPL as the target. Invasive fungal infections have increased in frequency and strength to prevail over the most potent antifungals available, and hence improved treatment options are urgently required. Though there has been much progress over the past decade in the development of antifungal therapies including echinocandins (eg caspofungin and micafungin) and derivatives of triazoles, voricanazole and prosaconazole, known as the second-generation triazoles, there are still unacceptably high levels of morbidity and mortality due to this infection (77). Agents with alternative modes of action are required that have activity against currently resistant strains. Already some proposed therapies include Nikkomycin Z, a peptide nucleoside inhibitor (78) and other peptides such as defensin, protegrins and histatins that penetrate into lipid bilayers to cause pores that lyse the cell to death. However, these therapies have risks of toxicity and development of resistance issues (79). For reasons analogous to those of studying SaBPL, CaBPL was characterised for the first time in this study. Much progress has been made in obtaining large quantities of this enzyme, for which crystallisation trials have begun in order to follow a structure-based drug design programme towards treatment of *Candida sp.* infections. In addition, an *in vitro* high-throughput screening assay has been designed to test small compound inhibitors of CaBPL. These methods will aid in the discovery of novel lead molecules that target an essential enzyme within *Candida* that differs from the drug targets already on the market, making resistance issues less likely.

The targeting of BPL may have a wide range of applications, not only to human health, but to other global issues such as plant pathogens (fungi being the largest group), which affect agricultural yields and represent a huge expense to farmers and consumers alike. Because of structural differences between fungal and plant BPLs, targeting BPLs from the pathogens of these resources may improve yields and quality of crops. There is a huge need and market, particularly in Australia, for novel inhibitors of pathogens affecting our agriculture. The development of inhibitors towards these pathogens BPLs may derive a pesticide that is both selective and potent and a better alternative to those already on the market.

Aside from using BPL as a potential drug target it is also important to fully understand the system and biological activity of this essential enzyme and why eukaryotic BPL's are approximately twice the size of prokaryotic BPLs. Through sequence homology searches and model building of *S. cerevisiae* BPL in comparison to the ligase domain of EcBPL, residues that are required for trapping the active biotin intermediate appear to be absent from the ligase domain of ScBPL. However, these may be supplied by the N-terminal domain, which was found to share homology with a domain of glutamine amidotransferase. A further development of this body of work would be to compare results of a similar SAXS experiment on HCS, and thus determine whether it to conforms to a similar predicted domain structure as ScBPL even though they only share approximately 19% sequence homology. It would then be a challenge to decipher what are the reasons for such variability in size and sequence to perform analogous tasks.

Chapter 8

General references

A collation of references for the General introduction (Chapters 1) and Discussion and future directions (chapter 7).

1. Walsh, C. (2003) *Nat Rev Microbiol* **1**, 65-70
2. Hiramatsu, K., Hanaki, H., Ino, T., Yabuta, K., Oguri, T., and Tenover, F. C. (1997) *J Antimicrob Chemother* **40**, 135-136
3. Palumbi, S. R. (2001) *Science* **293**, 1786-1790
4. Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007) *Nat Rev Drug Discov* **6**, 29-40
5. Ghannoum, M. A., and Rice, L. B. (1999) *Clin Microbiol Rev* **12**, 501-517
6. Hiemenz, J. W., and Walsh, T. J. (1996) *Clin Infect Dis* **22 Suppl 2**, S133-144
7. Patterson, T. F. (2005) *Lancet* **366**, 1013-1025
8. Poirier, J. M., and Cheymol, G. (1998) *Clin Pharmacokinet* **35**, 461-473
9. Manavathu, E. K., Cutright, J. L., and Chandrasekar, P. H. (1998) *Antimicrob Agents Chemother* **42**, 3018-3021
10. Johnson, L. B., and Kauffman, C. A. (2003) *Clin Infect Dis* **36**, 630-637
11. Keating, G. M. (2005) *Drugs* **65**, 1553-1567; discussion 1568-1559
12. Pasqualotto, A. C., and Denning, D. W. (2008) *J Antimicrob Chemother* **61 Suppl 1**, i19-30
13. Torres, H. A., Hachem, R. Y., Chemaly, R. F., Kontoyiannis, D. P., and Raad, II. (2005) *Lancet Infect Dis* **5**, 775-785
14. Mora-Duarte, J., Betts, R., Rotstein, C., Colombo, A. L., Thompson-Moya, L., Smietana, J., Lupinacci, R., Sable, C., Kartsonis, N., and Perfect, J. (2002) *N Engl J Med* **347**, 2020-2029
15. Balashov, S. V., Park, S., and Perlin, D. S. (2006) *Antimicrob Agents Chemother* **50**, 2058-2063
16. Park, S., and Perlin, D. S. (2005) *Microb Drug Resist* **11**, 232-238
17. Wakil, S. J., Stoops, J. K., and Joshi, V. C. (1983) *Annu Rev Biochem* **52**, 537-579
18. Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C., and Wood, H. G. (1988) *J. Biol. Chem.* **263**, 6461-6464
19. Knowles, J. R. (1989) *Ann. Rev. Biochem.* **58**
20. Artymiuk, P. J., Rice, D. W., Poirrette, A. R., and Willet, P. (1994) *Nat. Struct. Biol.* **1**, 758-760
21. Safro, M., and Mosyak, L. (1995) *Prot. Sci.* **4**, 2429-2432
22. Reche, P. A. (2000) *Protein Sci* **9**, 1922-1929
23. Kim, D. J., Kim, K. H., Lee, H. H., Lee, S. J., Ha, J. Y., Yoon, H. J., and Suh, S. W. (2005) *J Biol Chem* **280**, 38081-38089
24. Chapman-Smith, A., and Cronan, J. E., Jr. (1999) *Trends Biochem Sci* **24**, 359-363
25. Cronan, J. E., Jr. (1990) *J Biol Chem* **265**, 10327-10333
26. McAllister, H. C., and Coon, M. J. (1966) *J. Biol. Chem.* **241**, 2855-2861
27. Murtif, V. L., and Samols, D. (1987) *J Biol Chem* **262**, 11813-11816
28. Reed, K. E., and Cronan, J. E. J. (1991) *J. Biol. Chem.* **266**, 11425-11428
29. Chapman-Smith, A., Morris, T. W., Wallace, J. C., and Cronan, J. E., Jr. (1999) *J Biol Chem* **274**, 1449-1457
30. Denis, L., Grossemy, M., Douce, R., and Alban, C. (2002) *J Biol Chem* **277**, 10435-10444
31. Rodionov, D. A., Mironov, A. A., and Gelfand, M. S. (2002) *Genome Res* **12**, 1507-1516
32. Cronan, J. E., Jnr. (1989) *Cell* **58**, 427-429

33. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., and Matthews, B. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9257-9261
34. Chapman-Smith, A., Mulhern, T. D., Whelan, F., Cronan, J. E., Jr., and Wallace, J. C. (2001) *Protein Sci* **10**, 2608-2617
35. Weaver, L. H., Kwon, K., Beckett, D., and Matthews, B. W. (2001) *Protein Sci* **10**, 2618-2622
36. Kwon, K., and Beckett, D. (2000) *Protein Sci* **9**, 1530-1539
37. Brown, P. H., Cronan, J. E., Grotli, M., and Beckett, D. (2004) *J Mol Biol* **337**, 857-869
38. Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D., and Matthews, B. W. (2006) *J Mol Biol* **357**, 509-523
39. Otsuka, A. J. (1988) *J. Biol. Chem.*, 19577-19585
40. Eisenberg, M. A. (1973) *Adv. Enzymol.* **38**, 317-372
41. Abbott, J., and Beckett, D. (1993) *Biochemistry* **32**, 9649-9656
42. Prakash, o., Eisenberg, M, A. (1979) *Proc Natl Acad Sci U S A*, 5592-5595
43. Bagautdinov, B., Kuroishi, C., Sugahara, M., and Kunishima, N. (2005) *J Mol Biol* **353**, 322-333
44. Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2007) *Acta Crystallograph Sect F Struct Biol Cryst Commun* **63**, 334-337
45. Bagautdinov, B., Matsuura, Y., Bagautdinova, S., and Kunishima, N. (2008) *J Biol Chem*
46. Stolz, J., Ludwig, A., and Sauer, N. (1998) *FEBS Lett* **440**, 213-217
47. Hasslacher, M., Ivessa, A. S., Paltauf, F., and Kohlwein, S. D. (1993) *J Biol Chem* **268**, 10946-10952
48. Hoja, U., Marthol, S., Hofmann, J., Stegner, S., Schulz, R., Meier, S., Greiner, E., and Schweizer, E. (2004) *J Biol Chem* **279**, 21779-21786
49. Brewster, N. K., Val, D. L., Walker, M. E., and Wallace, J. C. (1994) *Arch. Biochem. Biophys.* **311**, 62-71
50. Genbauffe, F. S., and Cooper, T. G. (1986) *Mol. Cell. Biol.* **6**, 3954-3964
51. Kim, H. S., Hoja, U., Stolz, J., Sauer, G., and Schweizer, E. (2004) *J Biol Chem* **279**, 42445-42452
52. Lamhonwah, A.-M., Quan, F., and Gravel, R. A. (1987) *Arch. Biochem. Biophys.* **254**, 631-636
53. Perez-Vazquez, V., Uribe, S., and Velazquez-Arellano, A. (2005) *J Nutr Biochem* **16**, 438-440
54. Polyak, S. W., Chapman-Smith, A., Brautigan, P. J., and Wallace, J. C. (1999) *J Biol Chem* **274**, 32847-32854
55. Leon-Del-Rio, A., Leclerc, D., Akerman, B., Wakamatsu, N., and Gravel, R. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4626-4630
56. Cronan, J. E., Jr., and Wallace, J. C. (1995) *FEMS Microbiol Lett* **130**, 221-229
57. Campeau, E., and Gravel, R. A. (2001) *J Biol Chem* **276**, 12310-12316
58. Derewenda, z. s. (2004) *Methods*, 354-363
59. Sweet, C. W. C. a. R. M. (1997) *Macromolecular Crystallography*, Academic press, New York
60. C.W Carter Jr, R. M. S. (2003) *Macromolecular Crystallography*, Elsevier, Calafornia
61. Delano, W. L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto

62. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132
63. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng* **8**, 127-134
64. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) *Bioinformatics* **23**, 2947-2948
65. Thaver, D., Ali, S. A., and Zaidi, A. K. (2009) *Pediatr Infect Dis J* **28**, S19-21
66. Freiberg, C., Brunner, N. A., Schiffer, G., Lampe, T., Pohlmann, J., Brands, M., Raabe, M., Habich, D., and Ziegelbauer, K. (2004) *J Biol Chem* **279**, 26066-26073
67. Pendini, N. R., Polyak, S. W., Booker, G. W., Wallace, J. C., and Wilce, M. C. (2008) *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 520-523
68. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C., Wallace, J. C., and Polyak, S. W. (2008) *Biochim Biophys Acta* **1784**, 973-982
69. Mayende, L., Swift, R., Bailey, L., Pendini, N. R., Wallace, J. C., Polyak, S. W. a., and Booker, G. W. (2008) Domain structure of human holocarboxylase synthetase: Evidence of an interaction between the N-terminal and C-terminal halves., University of Adelaide
70. Fuchshuber, A., Suormala, T., Roth, B., Duran, M., Michalk, D., and Baumgartner, E. R. (1993) *Eur J Pediatr* **152**, 446-449
71. Bailey, L. M., Ivanov, R. A., Jitrapakdee, S., Wilson, C. J., Wallace, J. C., and Polyak, S. W. (2008) *Hum Mutat* **29**, E47-57
72. Qiagen. (1997) QIAexpress Detection Systems.
73. Pey, A. L., Ying, M., Cremades, N., Velazquez-Campoy, A., Scherer, T., Thony, B., Sancho, J., and Martinez, A. (2008) *J Clin Invest* **118**, 2858-2867
74. Balch, W. E., Morimoto, R. I., Dillin, A., and Kelly, J. W. (2008) *Science* **319**, 916-919
75. Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C., Wallace, J. C., and Polyak, S. W. (2008) *Arch Biochem Biophys* **479**, 163-169
76. Shao, P. L., Huang, L. M., and Hsueh, P. R. (2007) *Int J Antimicrob Agents* **30**, 487-495
77. Sable, C. A., Strohmaier, K. M., and Chodakewitz, J. A. (2008) *Annu Rev Med* **59**, 361-379
78. Ganesan, L. T., Manavathu, E. K., Cutright, J. L., Alangaden, G. J., and Chandrasekar, P. H. (2004) *Clin Microbiol Infect* **10**, 961-966
79. Lupetti, A., Danesi, R., van 't Wout, J. W., van Dissel, J. T., Senesi, S., and Nibbering, P. H. (2002) *Expert Opin Investig Drugs* **11**, 309-318