Biotin protein ligase from *Saccharomyces cerevisiae*

by

Steven William POLYAK, B.Sc.(Hons)

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

Department of Biochemistry
University of Adelaide
Adelaide, South Australia
March, 2000

TABLE OF CONTENTS

SUMMARY

Biotinylation *in vivo* is an extremely selective post-translational event where the enzyme biotin protein ligase (BPL) catalyses the covalent attachment of biotin to one specific and conserved lysine residue of biotin dependent enzymes. This event is essential for the synthesis of catalytically active biotin dependent enzymes, a family of enzymes found ubiquitously throughout the biological world. As these enzymes catalyse necessary metabolic reactions, *in vivo* biotinylation is a critical event for the survival of all organisms. The work presented in this thesis aims to investigate the physical and catalytic properties of a BPL, and understand further the interaction of this enzyme with its protein substrates.

Catalytically active biotin protein ligase from *Saccharomyces cerevisiae* [EC 6.3.4.15] was overexpressed in *Escherichia coli* and purified to near homogeneity in three steps. Kinetic analysis demonstrated that the substrates ATP, biotin and the biotin accepting protein bind in an ordered manner in the reaction mechanism. Treatment with any of three proteases of differing specificity *in vitro* revealed that the sequence between residues 240 and 260 was extremely sensitive to proteolysis, suggesting that it forms an exposed linker between an N-terminal 27 kDa domain and the C-terminal 50 kDa domain containing the active site. The protease susceptibility of this linker region was considerably reduced in the presence of ATP and biotin. A second protease-sensitive sequence, located in the presumptive catalytic site, was protected against digestion by the substrates. Expression of N-terminally truncated variants

of the yeast enzyme failed to complement *E. coli* strains defective in biotin protein ligase activity. *In vitro* assays performed with purified N-terminally truncated enzyme revealed that removal of the N-terminal domain reduced BPL activity by greater than 3500-fold. The data presented in this thesis indicate that both the N-terminal domain and the C-terminal domain containing the active site are necessary for complete catalytic function.

The biotin accepting lysine, present in a conserved Met-Lys-Met motif, resides in a structured domain that functions as the BPL substrate. Phage display coupled with a genetic selection to identify determinants of the biotin domain (yPC-104) of yeast pyruvate carboxylase 1 (residues 1075-1178) required for interaction with BPL was employed. Mutants isolated using this strategy were analysed by *in vivo* biotinylation assays performed at both 30^{o}C and 37^{o}C. The temperature sensitive substrates were reasoned to have structural mutations, leading to compromised conformations at the higher temperature. This interpretation was supplemented by molecular modelling of yPC-104, since these mutants mapped to residues involved in defining the structure of the biotin domain. In contrast, substitution of the Met residue N-terminal to the target lysine with either Val or Thr produced mutations that were temperature insensitive in the *in vivo* assay. In addition, these two mutant proteins and wildtype yPC-104 showed identical susceptibility to trypsin, suggesting that their amino acid substitutions did not induce structural changes. Kinetic analysis of enzymatic biotinylation using purified Met→Thr/Val mutant proteins with both yeast and *E. coli* BPLs revealed that these substitutions had a strong effect upon K_m values but not k_{cat}. The Met→Thr mutant was a poor substrate for both BPLs whereas the Met→Val substitution was a poor substrate for bacterial BPL but had only a

2-fold lower affinity for yeast BPL than the wildtype peptide. Our data suggest that substitution of Thr or Val for the Met N-terminal of the biotinyl-Lys results in genuine BPL interaction mutants.

STATEMENT OF ORIGINALITY

This thesis contains no material that has been accepted for the award of any degree or diploma by any other University. To the best of my knowledge it contains no material that has previously been published by any other person, except where due reference has been made in the text. I consent to this thesis, when deposited in the University library, being available for photocopying and loan.

29/3/2000

Steven Polyak

ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor, Dr John Wallace, for giving me the opportunity to further my university studies by undertaking both my honours and Ph.D. in his laboratory. His enthusiasm, knowledge and fruitful discussions allowed my work to constantly proceed in a positive fashion. I wish to thank Professor Peter Rathjen for permission to undertake this research project in the Department of Biochemistry, University of Adelaide. I am eternally grateful to Dr Anne Chapman-Smith for her supervision, technical assistance and, most importantly, friendship throughout my Ph.D.

Secondly I wish to thank those collaborators who contributed to this project. Dr Terry Mulhern, now at the Department of Biochemistry and Molecular Biology, University of Melbourne, performed the molecular modelling of the yeast biotin domain. Dr John Cronan, Departments of Microbiology and Biochemistry, University of Illinois, for the initial design of the project and critical reading of the manuscripts arising from this research. Dr Murray Stewart, MRC Laboratory of Molecular Biology, Cambridge, for enlightening and enthusiastic discussions.

I also wish to acknowledge the contribution of past and present members of Dr John Wallace's group from whom I learnt so much; Dr Dale Val, Dr Vicky Avery, Dr Göran Forsberg, Dr Michelle Walker, Dr Grant Booker and Mrs Sally Grovenor, Dr Brionny Forbes, and Dr Sarawut Jitrapakdee. Additionally, I wish to thank the members of the Wallace group who have made the laboratory a fun place to work; Melinda Lucic, Samantha Lien, Sara Mortellaro, Mark Nezic, Graham Hobba, Francine Carrick, 'Gap' Arpornsuwan, Peter Brautigan, Caroline Beekman, Kerrie McNeil, and Dennise Turner. I finally wish to thank the members of the Biochemistry Department, University of Adelaide, for their friendship, especially Gavin and Sharon Chapman.

It would be remiss of me not to mention those people from outside the science world who have supported me throughout my studies. To my wife to be, Adele, I am forever grateful for your understanding, love and support. I would like to thank my parents and sister for all their support. And finally to Shane who has given me the opportunity to beat him each week at tennis!

This thesis is dedicated to the memories of Simon Owen, Julianna and Steven Alexanda Polyak; *May we find a cure.*

PUBLICATIONS ARISING FROM THIS THESIS

Refereed journals

Polyak, S. W., Chapman-Smith, A., Brautigan, P. & Wallace, J. C. (1999) Biotin protein ligase from *Saccharomyces cerevisiae*: the N-terminal domain is required for full activity *J. Biol. Chem.* 274(46), 32847-32854

Polyak, S. W. Chapman-Smith, A., Morris, T., Cronan, J. E. Jr. & Wallace, J. C. (1999) Product phage display and colony screening of apo-biotinyl domain of yeast Pyruvate Carboxylase *manuscript submitted to J. Biol. Chem.*

Conference Abstracts

Polyak, S. W., Chapman-Smith, A. & Wallace, J. C. (1996) Using phage display for investigating post-translational modifications *Proc. Aus. Soc. Biochem. Mol. Biol.* pos. 3-234

Polyak, S. W., Chapman-Smith, A. & Wallace, J. C. (1997) Defining the biotin ligase recognition site on pyruvate carboxylase *Proc Lorne Conference on Protein Structure and Function* pos. C-69

Polyak, S. W., Chapman-Smith, A. & Wallace, J. C. (1997) Investigating the interaction between biotin ligase and biotin domains using phage display and an *in vivo* screen *Proc. Aus. Soc. Biochem. Mol. Biol.* pos. B1-55

Polyak, S. W., Chapman-Smith, A. & Wallace, J. C. (1998) The conserved methionine residues in biotin domains are required for recognition by biotin ligase *Proc Lorne Conference on Protein Structure and Function* pos. B-33

Polyak, S. W., Brautigan, P., Chapman-Smith, A. & Wallace, J. C. (1998) Characterisation and domain mapping of recombinant yeast biotin ligase *Proc. Aus. Soc. Biochem. Mol. Biol.* pos. 16

Polyak, S. W., Brautigan, P., Chapman-Smith, A. & Wallace, J. C. (1999) The N-terminal domain of yeast biotin ligase is essential for full enzyme activity *Proc. Lorne Conference on Protein Structure and Function* pos B49

Polyak, S. W., Brautigan, P., Chapman-Smith, A. & Wallace, J. C. (1999) The N-terminal domain of yeast biotin ligase is required for complete enzymatic activity *Proc. Aus. Soc. Biochem. Mol. Biol.* pos. 072

Chapman-Smith, A., Polyak, S. W. & Wallace, J. C. (1999) The residues around biocytin affect the efficiency of in vivo biotinylation *Proc. ASBMB Satellite Con. on Enzyme Structure and Mechanism*

ABBREVIATIONS

Abbreviations used throughout this thesis are in accordance with those described in the *Journal of Biological Chemistry*. additional abbreviations are shown below.

ACC	Acetyl CoA carboxylase
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
BCCP	Biotin carboxyl carrier protein
BCIG	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BPL	Biotin protein ligase
BSA	Bovine serum albumin
DNA	Deoxynucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
gIIIp	Gene III product
IPTG	Isopropyl β-D-thiogalactopyranoside
KAPA	7 keto, 8 amino – pelargonic acid
MCD	Multiple carboxylase deficiency
NiNTA	Nickel-nitriloacetic acid
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PC	Pyruvate carboxylase
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMSF	Phenylmethylsulphonyl fluoride
SDS	Sodium dodecyl sulfate
TC	Transcarboxylase
TEMED	N,N,N',N'-teramethylethylenediamine
TOCSY	Total correlation spectroscopy
Tris	Tris-(hydroxymethyl)-methylamine
Tween 20	Polyoxyethylene-sorbitan monolaurate
VHT	Vitamin H transpoter

Chapter 1

INTRODUCTION

INTRODUCTION

1.1 BIOTIN

In 1901, Wilders made the important observation that yeasts require a component present in both yeast and meat extract. This essential growth factor was extracted from egg yolk by Kogl and Tonnis, who named the molecule biotin (Kogl & Tonnis, 1936). It was also observed that raw egg in the diet was toxic to animals (Steinitz, 1898) and skin lesions caused by this diet could be cured by treatment with a heat stable factor from yeast or liver. This factor was termed vitamin H (*Haut*, German = skin) and was subsequently shown to be identical to biotin (du Vigneaud *et al.*, 1940; Gyorgy *et al.*, 1940). Later, it was discovered that avidin, a biotin-binding glycoprotein, was the toxic component of raw eggs (Eakin *et al.*, 1941). The structure of biotin was determined in 1942 (du Vigneaud *et al.*, 1942) and was confirmed by chemical synthesis (Harris *et al.*, 1943) and X-ray analysis (Traub, 1956).

During the 1950s it became evident that biotin is involved as a cofactor for a class of enzymes which catalyse a variety of carboxyl transfer reactions (Visser & Kellogg, 1978). Lynen demonstrated that the enzyme β-methylcrotonoyl-CoA carboxylase, which catalyses a step in the degradation of leucine, was capable of carboxylating free biotin (Lynen *et al.*, 1959). It soon became apparent that in the cell biotin functions as a cofactor only when bound to protein. Evidence for this was reported by Kosow and coworkers who showed that biotin was covalently attached to the ε-amino group of a specific lysine group of propionyl CoA carboxylase (Kosow *et al.*, 1962). The existence of this type of bond had

previously been proposed by Folkers after isolation of ε-biotinyl-lysine, or biocytin, from yeast extract (Wolf *et al.*, 1952). It was subsequently demonstrated for other biotin enzymes that the biotin prosthetic group is attached in the same way (reviewed Visser & Kellogg, 1978; Samols *et al.*, 1988). Further evidence for biotin dependency came from the observation that avidin could inactivate this family of enzymes (Green, 1975).

Biotin is a member of the water soluble B-complex group of vitamins which is synthesised by higher plants and most fungi and bacteria. Humans, other mammals and birds cannot synthesise biotin *de novo* and therefore must obtain this essential micronutrient from exogenous sources. In mammals this absorption process occurs in the small intestine. Other important sites that play a role in normal biotin physiology include the liver (the principle site of biotin utilisation), the kidney (the site of reabsorption of filtered biotin) and the placenta (the site of transport of biotin from the maternal circulation to the developing embryo) (Dyer & Said, 1997). Humans obtain biotin from material synthesised by intestinal microflora and from dietary sources (Wolf, 1995). Dietary biotin exists in both free and protein-bound forms (reviewed Dyer & Said, 1997; Friedrich, 1988; Said, 1999). Protein-bound biotin is digested by gastrointestinal proteases and peptidases to biocytin and biotin-containing short peptides (Wolf, 1995). Biotin present in these compounds is recycled through the action of serum biotinidase which releases free biotin (Hymes & Wolf, 1999; Wolf, 1995).

1.2 BIOTIN ENZYMES

Biotin enzymes are a family of enzymes ubiquitously found throughout nature. All members of this family utilise biotin as a mobile carboxyl carrier

(reviewed Wood & Barden, 1977; Samols *et al.*, 1988; Knowles, 1989). As shown in Figure 1.1, these enzymes catalyse reactions in two partial reactions carried out at separate sites, and the biotin prosthetic group acts as a flexible arm carrying the carboxyl group between the two sites. Biotin enzymes can be divided into three classes depending on the nature of the original donor and final carboxyl acceptor (Wood & Barden, 1977; Samols *et al.*, 1988). These are the carboxylases (Class I), decarboxylases (Class II) and transcarboxylases (Class III). All eukaryotic enzymes belong to Class I, while prokaryotes contain enzymes from all three classes (Figure 1.2).

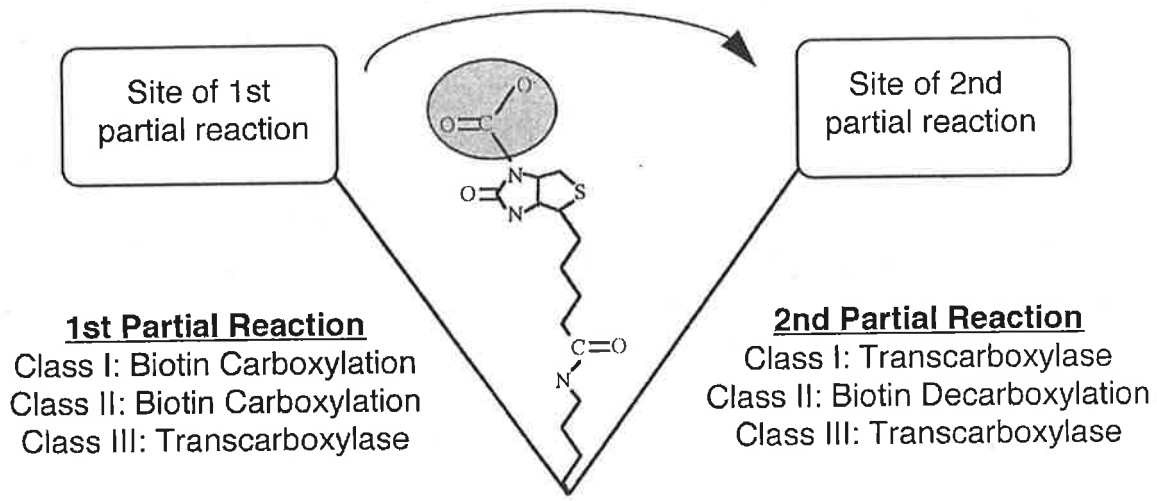


Figure 1.1: Biotin mediated carboxyl transfer between reaction subsites of biotin enzymes. The carboxylate anion, complexed to biotin, is shown in the shaded circle. Figure adapted from (Attwood, 1995).

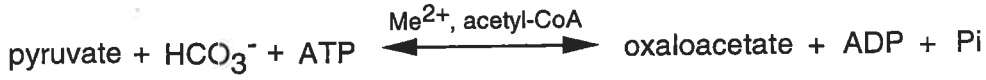
The mechanism of biotin carboxylation in the Class I enzymes has been extensively investigated (reviewed Wood & Barden, 1977; Attwood, 1995 Knowles, 1989). Experiments on propionyl CoA carboxylase (Kaziro *et al.*, 1962) and pyruvate carboxylase (Keech & Utter, 1963; Cooper & Wood, 1971) showed that the source of CO_2 in the reaction is bicarbonate and the reaction is dependent upon the presence of ATP (Eq. 1).

Figure 1.2: The reactions catalysed by the biotin dependent enzymes.

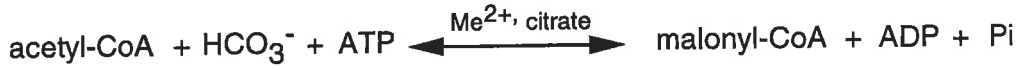
The reactions catalysed by the biotin-dependent enzymes are shown (Wood & Barden, 1977) along with the known cofactors. Me^{2+} represents a divalent metal ion, acetyl CoA is an allosteric activator of PC and citrate is an allosteric activator of acetyl CoA carboxylase (Lane *et al.*, 1974). Figure adapted from (Jitrapakdee, 1999).

Class I: CARBOXYLASES

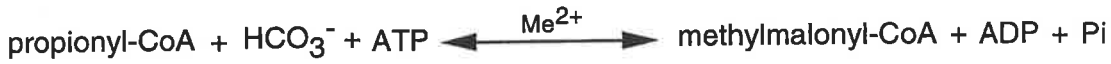
pyruvate carboxylase [EC 6.4.1.1]



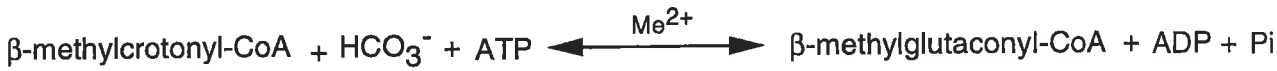
acetyl-CoA carboxylase [EC 6.4.1.2]



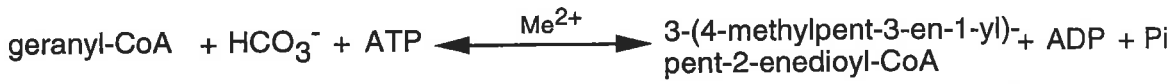
propionyl-CoA carboxylase [EC 6.4.1.3]



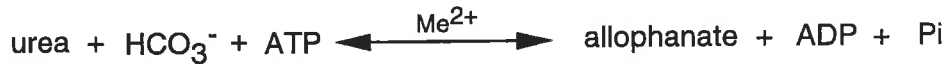
β-methylcrotonyl-CoA carboxylase [EC 6.4.1.4]



geranyl-CoA carboxylase [EC 6.4.1.5]



urea carboxylase [EC 6.3.4.6]



Class II: DECARBOXYLASES

oxaloacetate decarboxylase [4.1.1.3]

methylmalonyl-CoA decarboxylase [EC 4.1.1.41]

glutaconyl-CoA decarboxylase [EC 4.1.1.70]

Class III: TRANSCARBOXYLASE

transcarboxylase [2.1.3.1]

$$MgATP + HCO_3^- + \text{Enz-biotin} \rightleftharpoons MgADP + Pi + \text{Enz-biotin-}CO_2^- \quad \text{(Eq. 1)}$$

Biotin is carboxylated at the 1'-nitrogen probably after the formation of a very labile carboxyphosphate intermediate (Knowles, 1989; Phillips *et al.*, 1992). The carboxyl group is subsequently transferred to biotin.

The Class II decarboxylases differ from the Class I carboxylases in that the reactions are ATP-independent. The decarboxylases of anaerobic prokaryotes all catalyse decarboxylation of a specific β-keto acid or acyl-CoA coupled to sodium ion export against a concentration gradient (Dimroth, 1985). Thus, this class of enzyme serves as an important energy transducer that does not require ATP. There is only one known member belonging to the Class III biotin enzymes; *Propionibacterium shermanii* transcarboxylase (TC). TC catalyses the formation of methylmalonyl-CoA using oxaloacetate as a carboxyl donor (Samols *et al.*, 1988) and exists as a multimeric complex composed of 30 polypeptides of three different types; the 12S subunit involved in CoA ester binding, the 5S subunit required for keto-acid binding and a 1.3S biotin containing subunit.

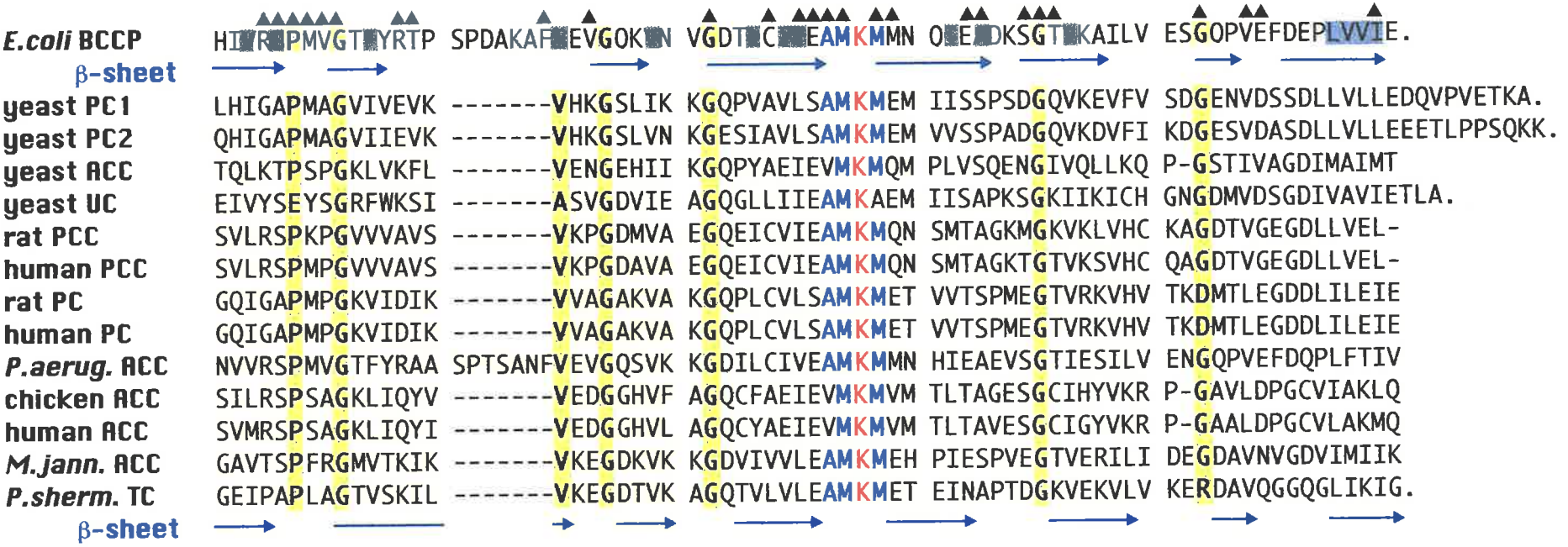
1.3 BIOTIN DOMAINS

The biotin prosthetic group is covalently attached via the ε-amino group of one specific lysine residue within the biotin enzymes (Samols *et al.*, 1988). This biotin-accepting lysine is found in an tetrapeptide sequence, Ala-Met-Lys-Met, which is extremely conserved amongst all biotin enzymes (Figure 1.3). Additionally, the primary structure surrounding the biocytin residue shows a high degree of homology between a wide range of enzymes and species (Figure 1.3)

Figure 1.3: Multiple sequence aligment of the biotinyl domains of various biotin carboxylases.

The biotinylated lysine residue (red) is present in the conserved Ala-Met-Lys-Met motif (blue). The highly conserved residues in the domain are highlighted (yellow). Residues forming the β-strands (arrows) in the 3-dimensional structures of BCCP (Atthappily & Hendrickson, 1995) and transcarboxylase (Reddy *et al.*, 1998) are shown under the sequences, as are the residues contributing to the hydrophobic core (blue shading). Positions at which amino acid substitution is known to reduce the efficiency of biotinylation are indicated above by ▲ (Murtif & Samols, 1987, Leon-Del Rio *et al.*, 1994, Chapman-Smith *et al.*, 1999, This study). The sequences shown were obtained from the following sources: *E. coli* BCCP (Li & Cronan, 1992), *S. cerevisiae* PC1 (Lim *et al.*, 1988), PC2 (Val *et al.*, 1995), ACC (Al-Feel *et al.*, 1992) and UC (Genbauffe & Cooper, 1991), rat PC (Jitrapakdee *et al.*, 1996) and PCC (Browner *et al.*, 1989), human PC (Walker *et al.*, 1995), PCC (Lamhonwah *et al.*, 1987) and ACC (Ha *et al.*, 1994), *Pseudomonas aerogenosa* ACC (Best & Knauf, 1993), chicken ACC (Takai *et al.*, 1988), *Methanococcus jannaaschii* ACC (Bult *et al.*, 1996) and *Propionibacterium shermanii* TC (Maloy *et al.*, 1979).

Abbreviations: BCCP; biotin carboxyl carrier protein; PC, pyruvate carboxylase; ACC acetyl CoA carboxylase; PCC, Propionyl CoA carboxylase; UC, urea carboxylase; TC, transcarboxylase.


E.coli BCCP
β-sheet
yeast PC1
yeast PC2
yeast ACC
yeast UC
rat PCC
human PCC
rat PC
human PC
P.aerug. ACC
chicken ACC
human ACC
M.jann. ACC
P.sherm. TC
β-sheet

(Samols *et al.*, 1988). Deletion studies on enzymes from both prokaryotes (Cronan, 1990; Li & Cronan, 1992; Stolz *et al.*, 1998) and eukaryotes (Leon-Del-Rio & Gravel, 1994; Val *et al.*, 1995) have shown that 30-35 amino acid residues either side of biocytin are necessary to specify biotinylation. Collectively these structures, able to incorporate biotin *in vivo*, are termed 'biotin domains'. The removal of determinants necessary to define the structure of a biotin domain by truncation or mutation results in a molecule which is unable to be biotinylated (Murtif & Samols, 1987; Cronan, 1990; Leon-Del-Rio & Gravel, 1994; Chapman-Smith *et al.*, 1999).

In the life cycle of a biotin domain, the molecule participates in a number of heterologous protein:protein interactions. First it serves as a substrate in the biotinylation reaction. Here the biotin prosthetic group is covalently attached to the biotin domain through the enzymatic action of biotin protein ligase (BPL). The attachment of biotin facilitates the binding of a carboxyl group at the first partial reaction site of the biotin enzyme. Finally the carboxyl group is transferred to the site of the second partial reaction. In the present study the interaction between a biotin domain and a BPL was investigated.

The formation of multimeric protein complexes, characteristic of all biotin enzymes, is not necessary for substrate recognition by BPLs. Wood and coworkers demonstrated that the biotin accepting subunit of the *P. shermanii* TC was biotinylated at the same rate as the enzyme complex (Wood *et al.*, 1980). A more comprehensive kinetic analysis of enzymatic biotinylation upon the biotin carboxyl carrier protein (BCCP) of *E. coli* acetyl CoA carboxylase was recently performed (Nenortas & Beckett, 1996). Comparison of the ability of full length protein and the C-terminal 87 amino acid residues of BCCP (BCCP-87) to function as substrates for *E. coli* BPL showed the enzyme had an equal affinity for

both proteins (Nenortas & Beckett, 1996). Therefore, the information required for association with BPL is present within the structured biotin domain.

BPLs from various sources have been found to recognise and biotinylate acceptor proteins from very different sources. McAllister and Coon (1966) first showed extracts containing BPLs from rabbit liver, yeast and the *P. shermanii* were able to activate enzyme substrates from rabbit and bacteria via the attachment of the biotin prosthetic group (McAllister & Coon, 1966). Evidence of cross species recognition *in vivo* was demonstrated when the 1.3S subunit of TC was recombinantly expressed in *E. coli* and shown to be a substrate for the bacterial BPL (Murtif *et al.*, 1985). Truncation analysis of the 1.3S TC subunit revealed that the minimum amount of information required to specify biotinylation was present in the 75 C-terminal amino acid residues (Cronan, 1990). This minimal peptide, fused to β-galactosidase, was biotinylated *in vivo* by the BPLs from *E. coli*. Similarly BCCP-87 was shown to be a stable protein which can function as a substrate for *E. coli* (Chapman-Smith *et al.*, 1994), insect (Lerner & Saiki, 1996) and human (Minich *et al.*, 1998) BPLs. Furthermore, the biotin domains of yeast pyruvate carboxylase 1 (yPC-1) (Val *et al.*, 1995) and human propionyl CoA carboxylase (Leon-Del-Rio & Gravel, 1994) have been shown to be substrates for bacterial BPL *in vivo*. Together these data suggest that all biotin domains fold into an essentially common tertiary structure which is recognised by all BPLs.

The use of biotin domains as fusion partners in the design of recombinant proteins has become a powerful tool in a number of biotechnological and molecular biology applications (Table 1.1). This strategy permits the specific biotinylation of proteins both *in vivo* and *in vitro* without chemical modification.

The presence of a biotin moiety in fusion proteins facilitates purification and detection of these proteins using well established avidin or streptavidin based techniques (Korpela, 1984; Wilchek & Bayer, 1990). The structured biotin domain usually does not interfere with the function of its protein partner and can be cleaved from the protein using site specific proteases if required.

Recently a 13 amino acid peptide was identified which functions as a substrate for *E. coli* BPL *in vivo* (Shatz, 1993). Kinetic analysis of this peptide revealed that the molecule is recognised by the bacterial BPL with an equal affinity as BCCP-87 (Beckett *et al.*, 1999), the native substrate for the enzyme. However, the peptide does not function as a substrate for either insect (Duffy *et al.*, 1998) or plant (Tissot *et al.*, 1998) BPLs. Investigation into the conformation of the peptide using NMR has shown that in solution the peptide is unstructured (Beckett personal comm). It seems likely that this peptide complexes with the bacterial BPL and reconforms to a structure which is complementary to the enzyme's active site.

Table 1.1: Applications of Recombinant Biotin Domain Fusion Proteins

Fusion protein	Host cell	Application	Reference
Proteins fused to *E. coli* BCCP-87			
MutS	*E. coli*	Isolation of DNA:protein complexes	Geschwind *et al.*, 1996
Retinoblastoma protein	*E. coli*	Isolation of protein:protein complexes	Wang *et al.*, 1996
Topoisomerase I	*S. frugiperda*	Isolation of DNA:protein complexes	Lerner & Saiki, 1996
Thyrotropin receptor	*H. sapien*	Screening for autoantibodies	Minich *et al.*, 1998
Maltose binding protein	*S. frugiperda*	Labelling protein with biotin	Duffy *et al.*, 1998
human Fab fragments	*E. coli*	Antibody screening	Sibler *et al.*, 1999
Proteins fused to *K. pneumoniae* OD biotin domain			
Citrate carrier	*K. pneumoniae*	Protein purification	Pos *et al.*, 1994
Plant sucrose carrier	*S. cerevisiae*	Purification from membranes	Stolz *et al.*, 1995
hexose-proton symporter	*S. pombe*	Protein purification	Caspari *et al.*, 1996
λ phage capsid protein D	*E. coli*	Phage display	Stolz *et al.*, 1998
Proteins fused to *S. cerevisiae* PC1 biotin domain			
Cin8p spindle motor	*S. cerevisiae*	Microtubule synthesis	Gheber *et al.*, 1999
M13 gIIIp coat protein	*E. coli*	Phage display	This study
Proteins fused to *P. shermanii* TC biotin domain			
β galactosidase	*E. coli* and *S. cerevisiae*	Mapping of minimal biotin domain	Cronan, 1990
Alkaline phosphatase & Maltose binding protein	*E. coli*	Topology of membrane proteins	Jander *et al.*, 1996

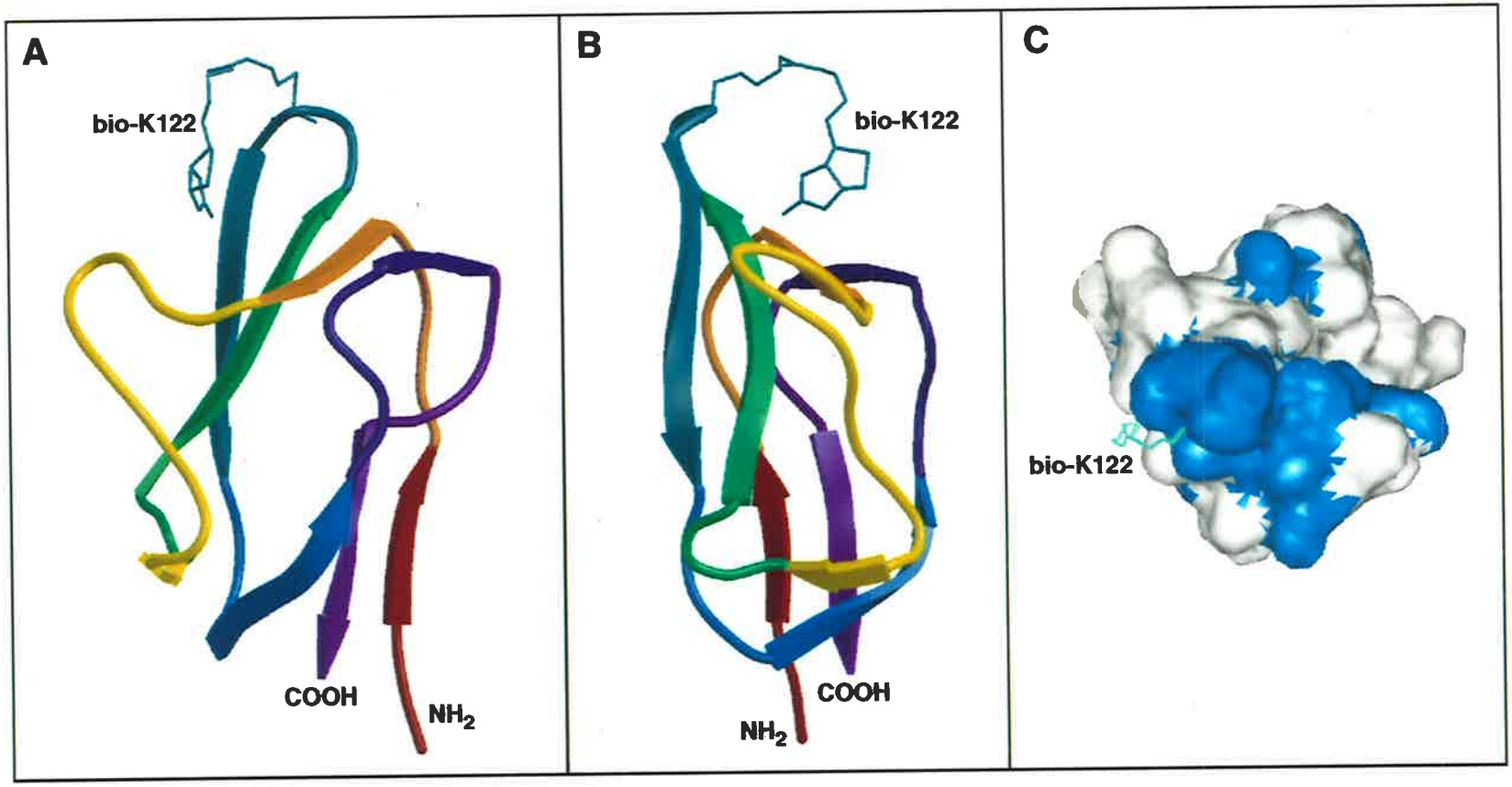
1.3.1 Biotin Domain Structure

The structures of two biotin domains have been determined; that of the *E. coli* BCCP-87 (Athappilly & Hendrickson, 1995; Roberts *et al.*, 1999) (Figure 1.4) and the 1.3S subunit of *P. shermanii* transcarboxylase (Reddy *et al.*, 1998). These proteins are structurally analogous to the lipoyl domains of 2-oxo acid dehydrogenase multienzyme complexes (Brocklehurst & Perham, 1993; Dardel *et al.*, 1993) which also undergo an analogous post-translational modification. These domains are a flattened β-barrel structure comprising two 4-stranded β-sheets with the N- and C- terminal residues close together at one end of the structure. At the other end of the molecule biotinyl- or lipoyl-lysine reside on a highly exposed, tight hairpin loop between β-strands four and five. The structure of the domain is stabilised by a core of hydrophobic residues which are important structural determinants. The conserved glycine residues found in these domains (Figure 1.3) occupy β-turns linking the β-strands.

BCCP-87 contains an additional seven amino acid residues, a feature which appears to be unique amongst all biotin domains with the exception of some prokaryotic acteyl CoA carboxylases (Best & Knauf, 1993). This region of peptide forms a 'thumb' structure between the β2 and β3 strands and contacts the biotin moiety in both the crystal (Athappilly & Hendrickson, 1995) and solution (Roberts *et al.*, 1999) structures. Determination of the solution structure of the 1.3S subunit of TC has revealed that no thumb structure is present in this biotin domain (Reddy *et al.*, 1998). Furthermore, the biotin moiety does not contact TC (Reddy *et al.*, 1997). The biological significance of this is not understood.

Figure 1.4: Structure of *E. coli* BCCP-87

The three dimensional structure of the 87 C-terminal amino acid residues of the biotin carboxyl carrier protein (BCCP) of *E. coli* acetyl CoA carboxylase has been determined by X-ray crystallography (Athappilly & Hendrickson, 1995) and NMR (Roberts *et al.*, 1999). Panel A shows the overall fold of the domain with the biotin moiety attached to lysine 122. The β-strands are coded from red at the N-terminus to purple at the C-terminus. The protruding thumb structure, shown here in yellow, most likely does not appear in other biotin domains. Panel B is the view of the molecule rotated by 90° around the Y-axis. Panel C is a space-filling representation in which the residues where substitutions decrease the efficiency of biotinylation are coloured blue (Murtif & Samols, 1987; Leon-Del-Rio & Gravel, 1994; Chapman-Smith *et al.*, 1999). This viewpoint is that from the upper right of the left panel. Figure reproduced from (Chapman-Smith & Cronan, 1999b)


A
bio-K122
COOH
NH2
B
bio-K122
COOH
NH2
C
bio-K122

Conformational changes accompany the conversion of apo BCCP-87 to its holo form (Chapman-Smith *et al.*, 1997). It has been demonstrated that the apo form of the protein is more susceptible to disulphide dimer formation and incorporation of the sulphydryl modifiers fluorescein-5-maleimide and 5,5'-dithiobis(2-nitrobenzoic acid) than the holo protein (Chapman-Smith *et al.*, 1997). The side chain of the single cysteine in the apo molecule, present in the β4-strand, must be sufficiently exposed to the solvent to facilitate thiol modification. Upon biotinylation this side chain becomes buried and forms part of the hydrophobic core (Athappilly & Hendrickson, 1995; Roberts *et al.*, 1999). Additionally, holo BCCP-87 is more resistant to proteolysis than apo BCCP-87 (Chapman-Smith *et al.*, 1997) further suggesting conformational changes accompany biotinylation. Recently the solution structure of apo BCCP-87 was determined (Yao *et al.*, 1997; Roberts *et al.*, 1999). The investigators concluded that the structures of apo and holo BCCP-87 are essentially identical, although the packing of side chains is more favourable in the holo domain (Roberts *et al.*, 1999; Yao *et al.*, 1999). These subtle alterations in protein conformation differences may function as a signal for the dissociation of the BPL-biotin domain complex upon completion of the biotinylation reaction.

1.4 BIOTIN DOMAIN : BIOTIN LIGASE INTERACTIONS

Several investigators have reported mutations to biotin domains which effect the efficiency of enzymatic biotinylation (Leon-Del-Rio & Gravel, 1994; Murtif & Samols, 1987). In these studies site directed mutagenesis was employed to introduce specific amino acid substitutions. The collective results of these studies, together with data from Chapman-Smith and coworkers (1999), and this

thesis, shows that interaction mutations can occur throughout the structured biotin domain (Figure 1.3). This most likely represents the importance of the biotin domain's structure for substrate recognition.

Recently a powerful *in vivo* screen in *E. coli* for the isolation of biotin domain mutants with a lowered affinity for BPL was reported (Chapman-Smith *et al.*, 1999). Here mutants of BCCP-87 were overexpressed in the biotin auxotroph bacterial strain TM21 and grown on media containing limited biotin (4.1 nM). High affinity substrates for biotin ligase were toxic to the host as these peptides reduced the cellular pool of free biotin by converting it into a non-available protein-bound form. The isolation of low affinity biotin ligase substrates was achieved as mutants which exhibited a decrease in affinity for the enzyme relieved the selective pressure applied to these cells. The result was biotinylation of essential acetyl CoA carboxylase and cell growth. Several BCCP-87 mutants were isolated using this method and the peptides purified. Analysis of the mutants using *in vitro* biotinylation assays, combined with the rates of tryptic digestion and sulphydryl modification to confirm structural integrity, revealed that two of the peptides obtained were genuine BPL interaction mutants which did not induce conformational changes in the protein. Interestingly, one of these mutations, E119K, was a substitution of the residue at the end of the β4 strand, two residues N-terminal of biocytin. Modelling of the mutations obtained from this work and other studies (Leon-Del-Rio & Gravel, 1994; Murtif & Samols, 1987) onto the BCCP-87 structure suggests that the face of the molecule containing the biotin accepting lysine includes determinants required for BPL interaction (Figure 1.4, panel C). This surface is likely to provide a protein scaffold upon which BPL can dock with the domain.

The precise positioning of the biotin accepting lysine residue in the structured biotin domain also appears to be important for substrate recognition by BPL. Movement of this residue by one position either side of its native position produces a peptide which is not biotinylated *in vivo* (Reche *et al.*, 1998). These biotin domain analogues were generated by replacing each of the methionine residues either side of the target lysine with a lysine and substituting the biotin-accepting lysine itself with an alanine. The precise positioning of the target lysine is also essential for lipoylation as analogous mutations made to the lipoyl domain of pyruvate dehydrogenase produced peptides which could not be lipoylated (Wallis & Perham, 1994).

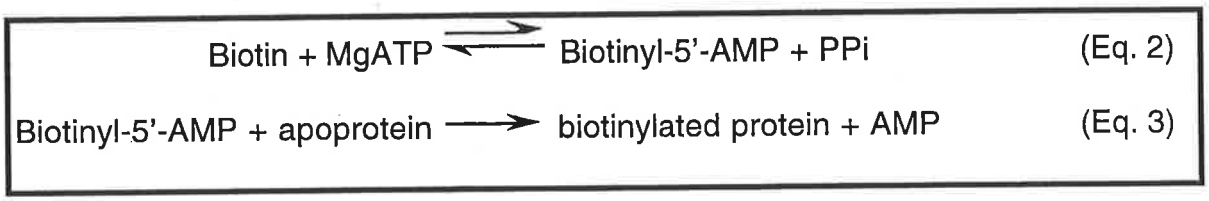
The properties of the amino acids residues immediately surrounding the target lysine are also important for substrate recognition. Conservative amino acid substitutions to the two methionine residues flanking the target lysine have revealed that these residues are not essential for the interaction with BPL (Shenoy & Wood, 1988; Leon-Del-Rio & Gravel, 1994). However, the methionine residues are important in biotin carboxylation of *P. shermanii* TC (Shenoy *et al.*, 1992) and a six amino acid residue peptide (Kondo *et al.*, 1984). The role of the residues flanking the target lysine in substrate recognition is not as important in lipoyl domains where a number of mutants with non-conservative substitutions were lipoylated (Wallis & Perham, 1994). Replacement of the Met-Lys-Met motif of *E. coli* BCCP with the Asp-Lys-Ala, motif characteristic of lipoyl domains, abolished biotinylation by the bacterial BPL (Reche & Perham, 1999). It is evident that the specificity of biotinylation *in vivo* requires the interaction of BPL with a structured protein substrate containing an exposed, precisely positioned lysine side chain.

Biotinylation and lipoylation *in vivo* by *E. coli* BPL and the two lipoyl attachment enzymes, LplA and LipB respectively, occur in a highly specific manner despite the striking structural similarity of the biotin and lipoyl accepting domains. In *E. coli* there is only one biotinylated protein, that is BCCP. This exists with the structurally homologous lipoyl domains from the 2-oxoglutarate dehydrogenase and pyruvate decarboxylase multienzyme complexes as well as the H-protein of the glycine cleavage system. The structure of BCCP-87 contains a protruding thumb between strands β2 and β3 (Athappilly & Hendrickson, 1995; Roberts *et al.*, 1999). This region of structure, not present in lipoyl domains, has been shown to inhibit lipoylation (Reche & Perham, 1999). Removal of the thumb produced a variant of BCCP-87 which was a substrate for both BirA and LplA *in vivo* and *in vitro* (Reche & Perham, 1999). In contrast, lipoyl domains contain a small surface loop between the β1 and β2 strands not present in biotin domains (Dardel *et al.*, 1993; Green *et al.*, 1995; Wallis *et al.*, 1996). This surface loop is in close proximity to the lipoyl moiety and has been shown to function in the reductive acylation of the prosthetic group (Wallis *et al.*, 1996). Removal of the loop does not affect lipoylation *in vivo* suggesting this structure does not contain determinants necessary for substrate recognition by the lipoylating enzymes (Wallis *et al.*, 1996). It is possible that this region of structure may play an inhibitory role in preventing biotinylation as when the lipoylation motif was replaced by the biotinylation motif in a lipoyl domain it could not be biotinylated *in vivo* (Wallis & Perham, 1994). These results imply that biotin and lipoyl ligases catalyse post-translational events of exceptional specificity.

1.5 BIOTIN LIGASE

1.5.1. Reaction Mechanism

All BPLs appear to catalyse biotinylation through the same reaction mechanism, shown in Figure 1.5 and below in equations 2 and 3;



In the first partial reaction, BPLs bind ATP and biotin in an ordered manner to form an enzyme-biotinyl-5'-AMP complex or holo BPL. At this stage the carboxyl group of inert biotin is activated by the addition of an adenylate group forming the intermediate biotinyl-5'-AMP. Siegel and coworkers first demonstrated that synthetic preparations of biotinyl-5'-AMP could be used to replace ATP and biotin in the reaction (Siegel *et al.*, 1965). This first partial reaction can be measured by detecting the formation of biotinyl-hydroxyamate from biotin, ATP, Mg^{2+} and hydroxylamine (Christner *et al.*, 1964; Morita *et al.*, 1998). In the second partial reaction holo BPL complexes with an apo biotin domain. Nucleophilic attack of the activated biotin by the amine group of the biotin-accepting lysine residue results in the transfer of biotin from the adenylate onto the biotin domain with AMP behaving as the leaving group (Xu & Beckett, 1997). This two step process is analogous to the tRNA synthetase catalysed linkage of amino acids to the 3' termini of tRNAs (Eisenberg, 1973; Fersht, 1985).

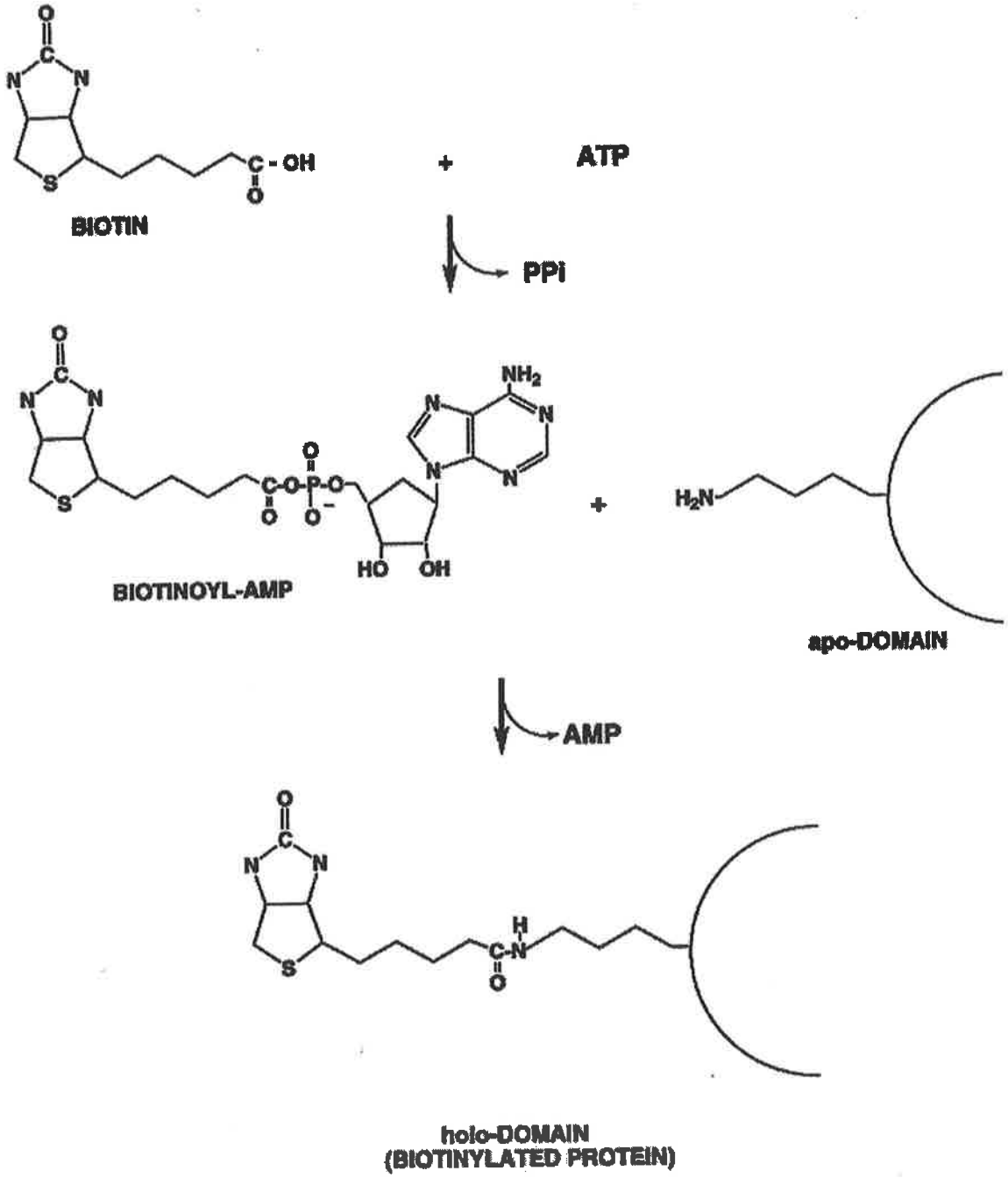


Figure 1.5: The reaction catalysed by biotin protein ligases

1.5.2. BirA

The most characterised and best understood BPL is the BirA protein from *E. coli* [E.C. 6.3.4.15]. Early genetic analysis of the *birA* locus by Barker and Campbell (Barker & Campbell, 1981a & b) revealed that the product encoded by this DNA was involved in repression of the biotin operon as well as containing holoacetyl CoA carboxylase synthetase activity. Isolation and sequencing of the *birA* gene (Barker & Campbell, 1981a; Howard *et al.*, 1985) revealed that the two functions were present in the one polypeptide sequence. The BPL protein (Cazzulo *et al.*, 1971; Lane *et al.*, 1964; Shenoy & Wood, 1988) and *bpl* gene (Bower *et al.*, 1995; Dombrosky *et al.*, 1994) have been isolated from several other prokaryotes, and an increasing number of *birA* homologues have been identified from genome sequencing projects in organisms as diverse as *Methanococcus jannaschii* (Bult *et al.*, 1996) and *Archaeoglobus fulgidus* (Klenk *et al.*, 1997). Similarity with *E. coli* BirA at the amino acid level suggests that all of these bacterial proteins are also bifunctional molecules.

1.5.2.1. Structure of BirA

The crystal structure of the 33.5 kDa protein has been determined at 2.3 Å resolution (Wilson *et al.*, 1992) (Figure 1.6). This revealed that BirA consists of three domains; An N-terminal DNA binding domain, a central catalytic domain and a small C-terminal domain of unknown function. The central catalytic domain contacts biotin and contains the nucleotide triphosphate binding motif GRGRRG. No function has been assigned to the small C-terminal domain.

The structure of BirA is an assembly of recognised protein folds. Noble and coworkers (Noble *et al.*, 1993) first noted that the C-terminal domain of BirA

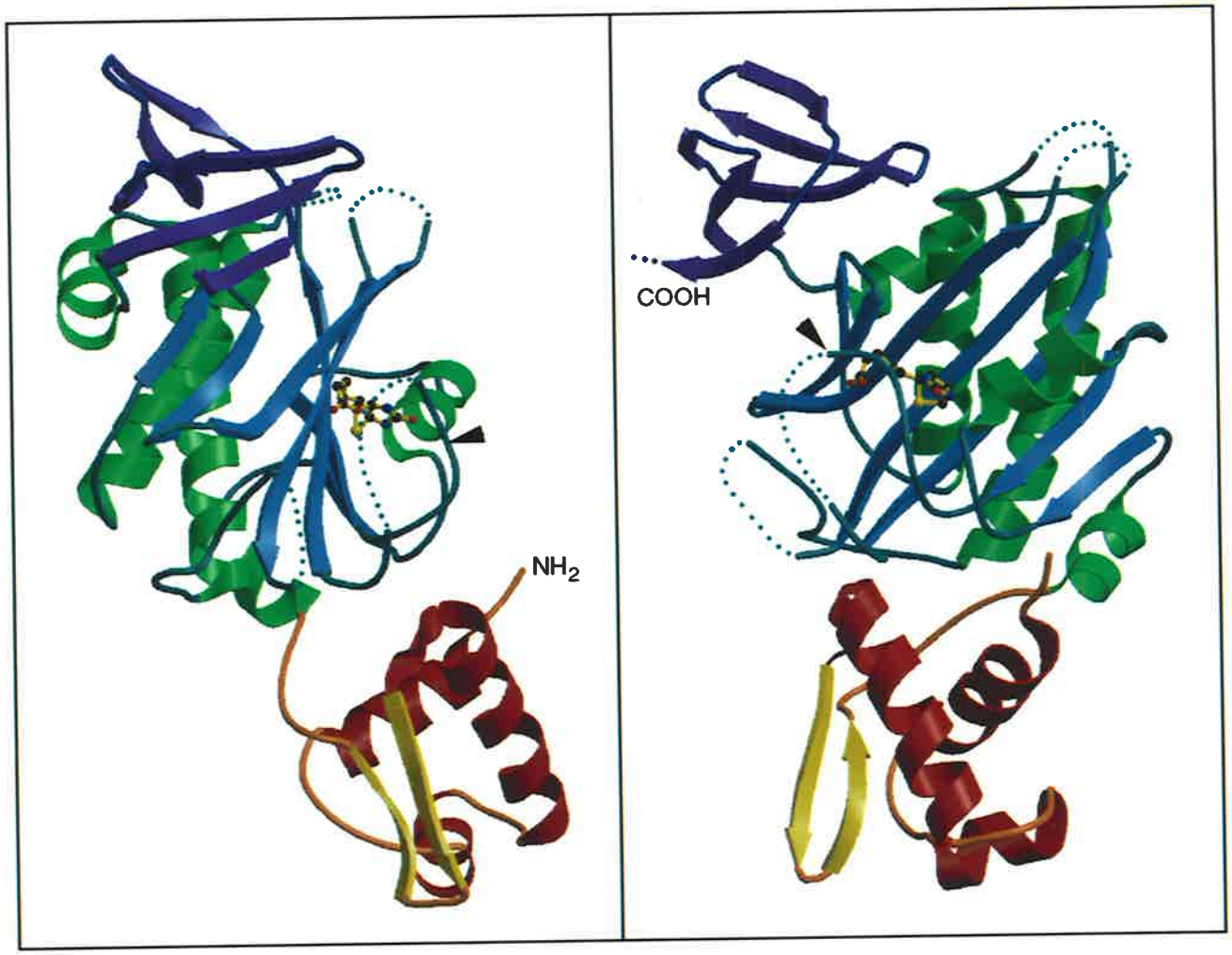
contains structural similarity to the Src homology 3 domain (SH3) of Fyn, a Src family tyrosine kinase. Using a protein-structure comparison algorithm, Russell and Barton observed structural similarity between the central domain of BirA and the phosphate binding SH2 domain of chicken Src (Russell & Barton, 1993). This domain in BirA is similarly involved in phosphate binding by forming biotinyl-5'-AMP from ATP and biotin. Thus the catalytic and C-terminal domains of BirA are a fusion of SH2 and SH3 domains. Additionally, the catalytic region of BirA is structurally analogous with the corresponding region of aminoacyl transfer-RNA synthetases (Artymiuk *et al.*, 1994). Eisenberg first reported that the reaction mechanisms of these two enzymes are similar as both function through an acyl-adenylate intermediate (Eisenberg, 1973). The sequence similarity between the two proteins is low making it impossible to show if the two proteins evolved from a single common ancestor inspite of their common structures.

1.5.2.2 Biology of BirA

BirA is a bifunctional protein. In addition to biotin ligase activity, the molecule is also the repressor of the biotin biosynthetic operon. This operon consists of five structural genes with the operator region, *bioO*, located between two clusters (Beckett & Matthews, 1997). The *bioO* sequence is a 40 base pair imperfect palindrome which represses initiation of transcription from two promoters, P_A and P_B, when complexed with BirA. This sequence contains two half sites which allows two BirA molecules to occupy the DNA (Abbott & Beckett, 1993; Streaker & Beckett, 1998a). The association between the two BirA molecules and the two operator half sites is a co-operative event as BirA is monomeric in solution at concentrations required to bind DNA (Abbott & Beckett,

Figure 1.6: Structure of BPL form *E. coli*, the BirA protein.

The structure of BirA was determined using X-ray crystallography (Wilson *et al.*, 1992). The molecule contains three domains. The N-terminal domain, represented by red α-helices and yellow β-sheets, has a helix-loop-helix fold for binding DNA. The central domain shows structural homology to SH2 domains and is represented by green α-helices and blue β-sheets. This domain is responsible for binding ATP and biotin. The structure shows biotin bound at the active site, with the nearby loop containing the ATP binding site indicated by an arrow. The C-terminal domain, with structural homology with SH3 domains, is represented by purple β-sheets. The right hand panel shows the molecule rotated by 90° around the Y-axis. The broken lines indicate the poorly defined regions of the molecule not seen in the electron density maps (Wilson *et al.*, 1992). Figure reproduced from (Chapman-Smith & Cronan, 1999b).


NH2
COOH

1993; Streaker & Beckett, 1998b). The interaction between BirA and DNA is facilitated through the N-terminal domain of the repressor. This α-helical domain contains a helix-turn-helix motif with a central DNA recognition helix required for binding to a specific DNA sequence (Wilson *et al.*, 1992).

BirA is a unique transcriptional repressor, in that it catalyses the synthesis of its own corepressor, biotinyl-5'-AMP. The holoBirA complex is the DNA binding form of the enzyme. Upon binding biotin, BirA undergoes a large conformational change which facilitates the binding of ATP (Xu *et al.*, 1995; Xu *et al.*, 1996). These conformational changes are thought to change the relative positions of the structural domains permitting both DNA binding and dimerisation to occur (Wilson *et al.*, 1992). The kinetics of the formation of the adenylated intermediate by BirA have been quantitatively analysed (Xu & Beckett, 1994). The holoBirA complex is quite stable (Xu & Beckett, 1994) and is proposed to be the most abundant enzyme form in the cell. The order of substrate binding in *E. coli* is believed to allow more responsive regulation of biotin biosynthesis. When the cellular demand for biotin is low, the holoBirA complex occupies the *bio* operator sequence and represses transcription of the biotin biosynthetic operon. As apo BCCP levels increase, biotin is transferred from the enzyme-bound adenylate to the protein bound form, with concomitant derepression of the biotin operon (Cronan, 1990). Kinetic analysis of the interaction of BirA with biotin indicates that formation of the repressor complex is highly sensitive to biotin (Beckett & Matthews, 1997; Chapman-Smith *et al.*, 1999; Xu & Beckett, 1994).

The N-terminal DNA binding domain of BirA appears to affect the affinity of the catalytic domain for both biotin and biotinyl-5'-AMP. A truncation mutant in which the N-terminal domain was absent was still able to catalyse biotin transfer,

but displayed a 100-fold decrease in the affinity for biotin and a 1000-fold decrease in the affinity for biotinyl-5'-AMP (Xu & Beckett, 1996). Biotin binding to this truncation caused no quenching of intrinsic protein fluorescence, as opposed to the 15% quenching observed with the intact enzyme (Xu & Beckett, 1996). These data suggest that quenching of fluorescence may be the result of the conformational changes which are induced by ligand binding and the truncated enzyme is compromised in its ability to go through these changes.

1.5.3. Eukaryotic BPLs

Unlike BirA, little structural information exists regarding eukaryotic members of the BPL family. The primary structure of human BPL has been reported by two groups. Suzuki and coworkers isolated a cDNA clone containing the human coding region by using primers derived from the amino acid sequence of purified bovine BPL (Suzuki *et al.*, 1994). Using an alternative approach, Leon-Del-Rio and coworkers obtained a similar cDNA clone using a complementation screen in a conditionally lethal strain of *E. coli* defective in BirA activity (Leon-Del-Rio *et al.*, 1995). The open reading frame in the human cDNA encodes a protein of 726 amino acid residues (predicted Mr 80.8 kD). Three possible initiation of transcription sites exist in the 5' region of the gene (Suzuki *et al.*, 1994; Leon-Del-Rio *et al.*, 1995). Three isoforms of human BPL, differing in their N-terminal sequence, were identified using an *in vitro* translation system and mRNA isolated from human placenta (Hiratsuka *et al.*, 1998). These studies suggest that all three initiation sites can be utilised by alternative splicing but neither the mechanisms for this nor the consequences of it have yet been determined.

A similar functional complementation approach was employed to isolate the *bpl* coding regions from *Saccharomyces cerevisiae* (Cronan & Wallace, 1995) and *Aradopsis thaliana* (Tissot *et al.*, 1997). These genes encode proteins of 690 (Mr 76.4 kDa) and 367 (Mr 41.1 kDa) amino acid residues respectively. In addition, a putative BPL sequence from *Schizosaccharomyces pombe* has been submitted in the EMBL database (Wood *et al.*, 1997) which encodes a 631 amino acid residue protein (Mr 70.6 kDa).

Sequence alignments of all BPLs reveal that the catalytic regions are highly conserved (Figure 1.7). Those residues in BirA which contact biotin have been determined in the crystal structure (Wilson *et al.*, 1992) and are invariant in all BPLs (Chapman-Smith & Cronan, 1999a). In addition, the GRGRRG motif, proposed to be associated with ATP binding, is also highly conserved. A conserved KWPND motif is present in BPLs and avidin but the biological significance of this sequence is unknown. However the primary structure outside of the catalytic domain is different between species (Figure 1.8). The N-terminal, DNA binding domain found in the prokaryotic BPLs does not exist in the higher organisms. Unlike bacteria, humans and *S. cerevisiae* are biotin prototrophs and therefore have no need to regulate biotin biosynthesis. Plants do synthesise biotin but the high biotin concentration found in their cells implies the absence of a strong repression mechanism. The N-terminal regions of the BPLs from higher organisms contain additional sequences which yield enzymes up to 2-fold larger than BirA (Figure 1.9). The structure and function of these additional residues is unknown.

Figure 1.7: Homology within the catalytic domains of biotin protein ligases.

Sequences for BPL proteins from *Escherichia coli* (Howard *et al.*, 1985), *Bacillus subtilis* (Bower *et al.*, 1995), *Arabidopsis thaliana* (Tissot *et al.*, 1997), *Saccharomyces cerevisiae* (Cronan & Wallace, 1995), *Homo sapiens* (Suzuki *et al.*, 1994; Dupuis *et al.*, 1996) and putative BPLs from *Paracoccus denitrificans* (Xu *et al.*, 1993), *Schizosaccharomyces pombe* (Wood *et al.*, 1997) and *Methanococcus janaschii* (Bult *et al.*, 1996) were aligned over their entire length using ClustalW. Residues making contact with biotin in the crystal structure of *E. coli* BirA (Wilson *et al.*, 1992) are indicated with ■ and the extent of the catalytic domain is marked using arrowheads. Amino acid changes in the catalytic domain caused by mutations affecting BPL activity in *E. coli* BirA are shown by ▲. Substitutions resulting in an increased K_m for biotin are in green and those affecting repressor function without altering biotin requirement, ie, suggesting residues involved in ATP binding or catalysis, in blue. Point mutations which give rise to the human condition MCD are indicated under the sequence by ●. Highly conserved residues are shaded, highlighting both the ATP binding motif (blue) and the KWPND sequence also present in avidin (red). Figure adapted from (Chapman-Smith & Cronan, 1999a).

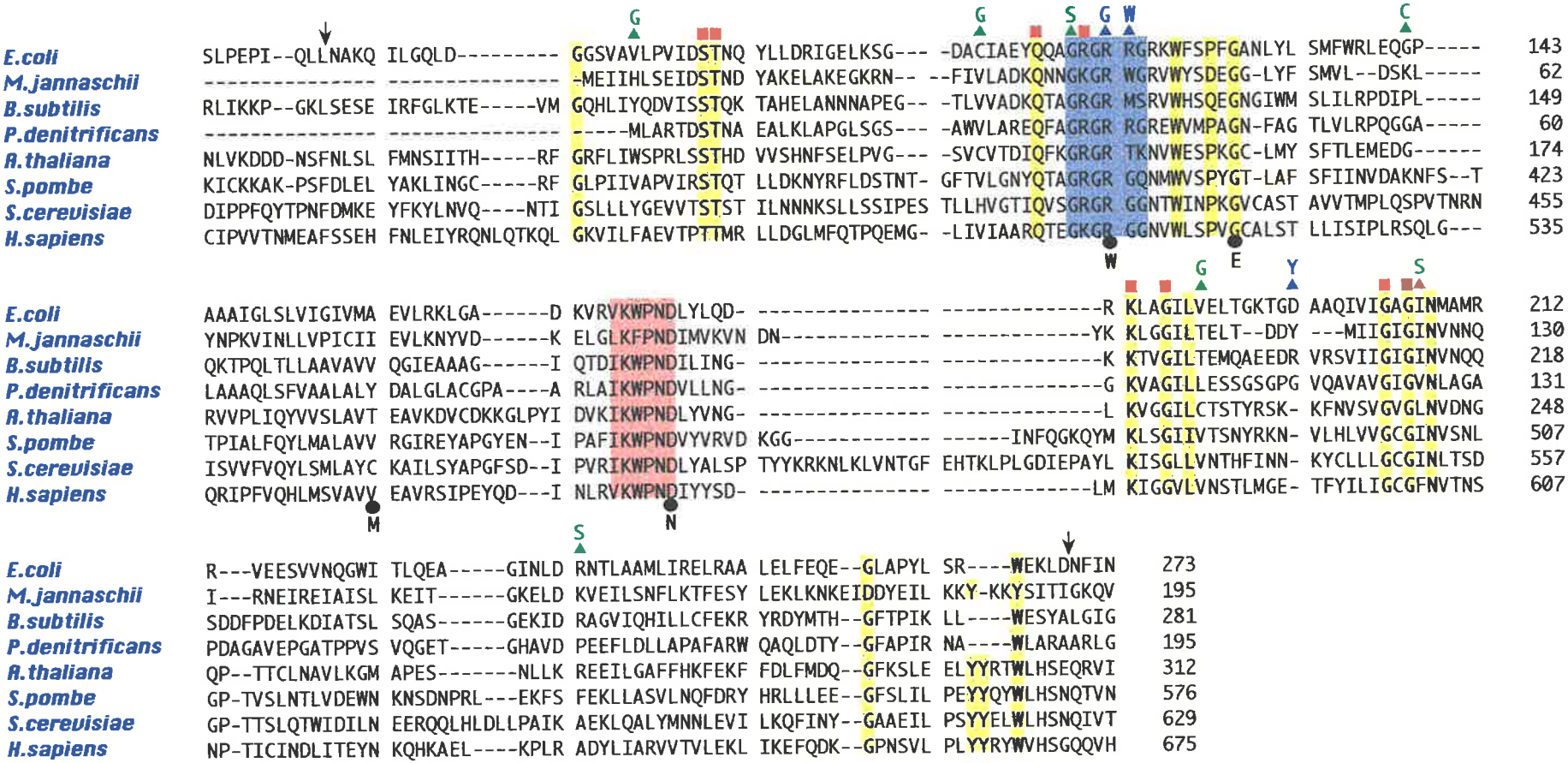
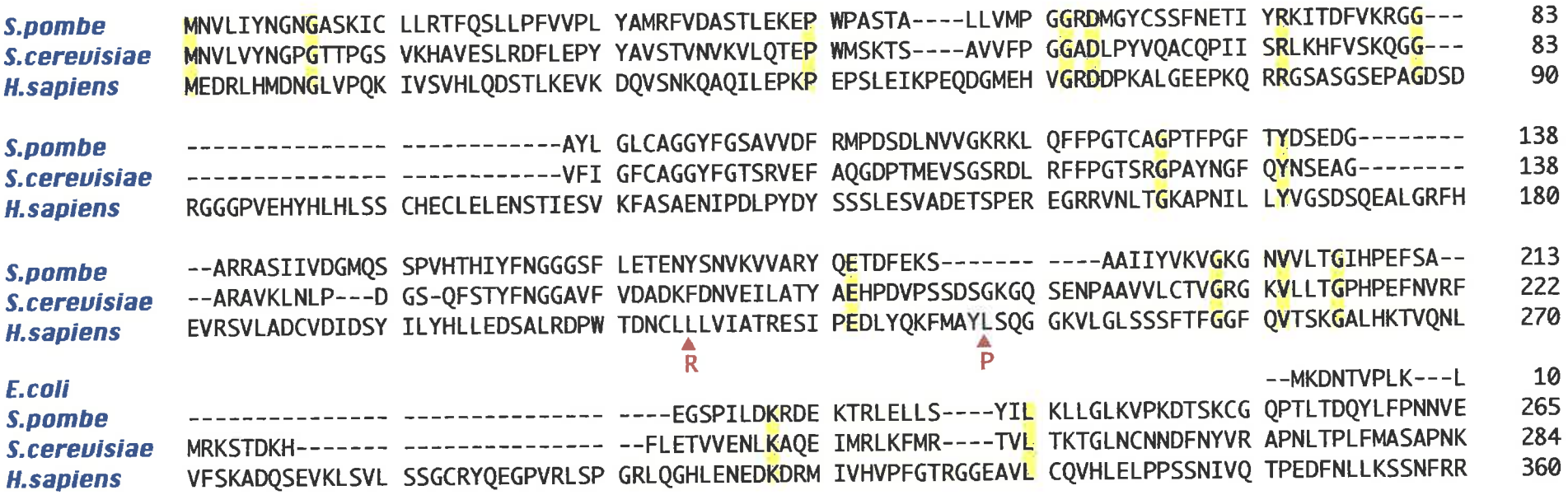

E.coli SLPEPI--QLLNAKQ ILGQLD--------- GGSVAVLPVIDSTNQ YLLDRIGELKSG--- -DACIAEYQQAGRGR RGRKWFSPFGANLYL SMFWRLEQGP----- 143
M.jannaschii ---------- ---------- -MEIIHLSEIDSTND YAKELAKEGKRN-- --FIVLADKQNNGKGR WGRVWYSDEGG-LYF SMVL--DSKL----- 62
B.subtilis RLIKKP--GKLSESE IRFGLKTE-----VM GQHLIYQDVISSTQK TAHELANNNAPEG-- -TLVVADKQTAGRGR MSRVWHSQEGNGIWM SLILRPDIPL----- 149
P.denitrificans ---------- ---------- -----MLARTDSTNA EALKLAPGLSGS--- -AWVLAREQFAGRGR RGREWVMPAGN-FAG TLVLRPQGGA----- 60
A.thaliana NLVKDDD-NSFNLSL FMNSIITH-----RF GRFLIWSPRLSSTHD VVSHNFSELPVG--- -SVCVTDIQFKGRGR TKNVWESPKGC-LMY SFTLEMEDG------ 174
S.pombe KICKKAK-PSFDLEL YAKLINGC-----RF GLPIIVAPVIRSTQT LLDKNYRFLDSTNT- GFTVLGNYQTAGRGR GQNMWVSPYGT-LAF SFIINVDAKNFS--T 423
S.cerevisiae DIPPFQYTPNFDMKE YFKYLNVQ----NTI GSLLLYGEVVTSTST ILNNNKSLLSSIPES TLLHVGTIQVSGRGR GGNTWINPKGVCAST AVVTMPLQSPVTNRN 455
H.sapiens CIPVVTNMEAFSSEH FNLEIYRQNLQTKQL GKVILFAEVTPTTMR LLDGLMFQTPQEMG- -LIVIAARQTEGKGR GGNVWLSPVGCALST LLISIPLRSQLG--- 535
E.coli AAAIGLSLVIGIVMA EVLRKLGA------D KVRVKWPNDLYLQD- ---------- ----------R KLAGILVELTGKTGD AAQIVIGAGINMAMR 212
M.jannaschii YNPKVINLLVPICII EVLKNYVD------K ELGLKFPNDIMVKVN DN---------- ----------YK KLGGILTELT--DDY ---MIIGIGINVNNQ 130
B.subtilis QKTPQLTLLAAVAVV QGIEAAAG------I QTDIKWPNDILING- ---------- ----------K KTVGILTEMQAEEDR VRSVIIGIGINVNQQ 218
P.denitrificans LAAAQLSFVAALALY DALGLACGPA----A RLAIKWPNDVLLNG- ---------- ----------G KVAGILLESSGSGPG VQAVAVGIGVNLAGA 131
A.thaliana RVVPLIQYVVSLAVT EAVKDVCDKKGLPYI DVKIKWPNDLYVNG- ---------- ----------L KVGGILCTSTYRSK- KFNVSVGVGLNVDNG 248
S.pombe TPIALFQYLMALAVV RGIREYAPGYEN--I PAFIKWPNDVYVRVD KGG---------- ------INFQGKQYM KLSGIIVTSNYRKN- VLHLVVGCGINVSNL 507
S.cerevisiae ISVVFVQYLSMLAYC KAILSYAPGFSD--I PVRIKWPNDLYALSP TYYKRKNLKLVNTGF EHTKLPLGDIEPAYL KISGLLVNTHFINN- KYCLLLGCGINLTSD 557
H.sapiens QRIPFVQHLMSVAVV EAVRSIPEYQD---I NLRVKWPNDIYYSD- ---------- ----------LM KIGGVLVNSTLMGE- TFYILIGCGFNVTNS 607
E.coli R---VEESVVNQGWI TLQEA-----GINLD RNTLAAMLIRELRAA LELFEQE--GLAPYL SR----WEKLDNFIN 273
M.jannaschii I---RNEIREIAISL KEIT------GKELD KVEILSNFLKTFESY LEKLKNKEIDDYEIL KKY-KKYSITIGKQV 195
B.subtilis SDDFPDELKDIATSL SQAS------GEKID RAGVIQHILLCFEKR YRDYMTH--GFTPIK LL----WESYALGIG 281
P.denitrificans PDAGAVEPGATPPVS VQGET-----GHAVD PEEFLDLLAPAFARW QAQLDTY--GFAPIR NA----WLARAARLG 195
A.thaliana QP--TTCLNAVLKGM APES-------NLLK REEILGAFFHKFEKF FDLFMDQ--GFKSLE ELYYRTWLHSEQRVI 312
S.pombe GP-TVSLNTLVDEWN KNSDNPRL---EKFS FEKLLASVLNQFDRY HRLLLEE--GFSLIL PEYYQYWLHSNQTVN 576
S.cerevisiae GP-TTSLQTWIDILN EERQQLHLDLLPAIK AEKLQALYMNNLEVI LKQFINY--GAAEIL PSYYELWLHSNQIVT 629
H.sapiens NP-TICINDLITEYN KQHKAEL----KPLR ADYLIARVVTVLEKL IKEFQDK--GPNSVL PLYYRYWVHSGQQVH 675

Figure 1.8 Multiple sequence alignment in the N-terminal regions of BPLs

Sequences for BPL proteins from *Escherichia coli* (Howard *et al.*, 1985), *Saccharomyces cerevisiae* (Cronan and Wallace, 1995), *Homo sapiens* (Suzuki *et al.*, 1994; Dupuis *et al.*, 1996) and putative *Schizosaccharomyces pombe* BPL (Wood *et al.*, 1997) were aligned over their entire length using ClustalW. The residues conserved between the three proteins are shaded yellow. Two positions of point mutations giving rise to the human condition of multiple carboxylase deficiency are shown in red (Suzuki *et al.*, 1994; Dupuis *et al.*, 1996). Figure reproduced from Chapman-Smith & Cronan (1999).


S.pombe MNVLIYNGNGASKIC LLRTFQSLLPFVVPL YAMRFVDASTLEKEP WPASTA----LLVMP GGRDMGYCSSFNETI YRKITDFVKRGG--- 83
S.cerevisiae MNVLVYNGPGTTPGS VKHAVESLRDFLEPY YAVSTVNVKVLQTEP WMSKTS----AVVFP GGADLPYVQACQPII SRLKHFVSKQGG--- 83
H.sapiens MEDRLHMDNGLVPQK IVSVHLQDSTLKEVK DQVSNKQAQILEPKP EPSLEIKPEQDGMEH VGRDDPKALGEEPKQ RRGSASGSEPAGDSD 90
S.pombe ---------- ----------AYL GLCAGGYFGSAVVDF RMPDSDLNVVGKRKL QFFPGTCAGPTFPGF TYDSEDG-------- 138
S.cerevisiae ---------- ----------VFI GFCAGGYFGTSRVEF AQGDPTMEVSGSRDL RFFPGTSRGPAYNGF QYNSEAG-------- 138
H.sapiens RGGGPVEHYHLHLSS CHECLELENSTIESV KFASAENIPDLPYDY SSSLESVADETSPER EGRRVNLTGKAPNIL LYVGSDSQEALGRFH 180
S.pombe --ARRASIIVDGMQS SPVHTHIYFNGGGSF LETENYSNVKVVARY QETDFEKS------- ----AAIIYVKVGKG NVVLTGIHPEFSA-- 213
S.cerevisiae --ARAVKLNLP---D GS-QFSTYFNGGAVF VDADKFDNVEILATY AEHPDVPSSDSGKGQ SENPAAVVLCTVGRG KVLLTGPHPEFNVRF 222
H.sapiens EVRSVLADCVDIDSY ILYHLLEDSALRDPW TDNCLLLVIATRESI PEDLYQKFMAYLSQG GKVLGLSSSFTFGGF QVTSKGALHKTVQNL 270
R
P
--MKDNTVPLK---L 10
E.coli
S.pombe ---------- ---------- ----EGSPILDKRDE KTRLELLS----YIL KLLGLKVPKDTSKCG QPTLTDQYLFPNNVE 265
S.cerevisiae MRKSTDKH------- ---------- --FLETVVENLKAQE IMRLKFMR----TVL TKTGLNCNNDFNYVR APNLTPLFMASAPNK 284
H.sapiens VFSKADQSEVKLSVL SSGCRYQEGPVRLSP GRLQGHLENEDKDRM IVHVPFGTRGGEAVL CQVHLELPPSSNIVQ TPEDFNLLKSSNFRR 360

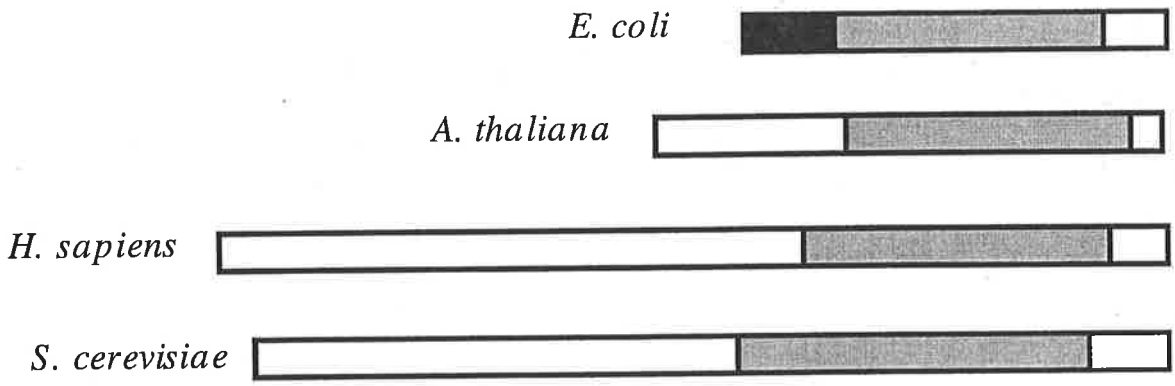


Figure 1.9: Relative sizes of BPLs. Schematic diagram showing the relative sizes of BPLs from bacteria, plant, humans and yeast. The shaded regions indicate the predicted catalytic regions of each of the molecules (Chapman-Smith & Cronan, 1999a). The solid box shows the N-terminal DNA binding domain of BirA.

1.5.4 Human BPL and disease

Human BPL is responsible for the biotinylation of four enzymes; acetyl CoA carboxylase (ACC), pyruvate carboxylase (PC), propionyl CoA carboxylase and β-methylcrotonyl CoA carboxylase. All of these enzymes are located in the mitochondrion with the exception of acetyl CoA carboxylase which is cytosolic. It appears that biotinylation can occur in either cellular compartment as BPL activity has been detected associated with the mitochondrion of other mammals such as rats (Cohen *et al.*, 1985), bovine liver (Chiba *et al.*, 1994) and a mouse cell line (Chang & Cohen, 1983). However the presence of human BPL in the mitochondrion and the existence of a suitable targeting signal in the coding region have not been confirmed.

The reactions catalysed by the four biotin enzymes are shown in Figure 1.2 The synthesis of malonyl CoA in the reaction performed by ACC is the first committed step in the synthesis of fatty acids. PC catalyses the generation of

oxaloacetate, a precursor for the synthesis of glucose and fat as well as some amino acids and neurotransmitters (Jitrapakdee & Wallace, 1999; Wallace *et al.*, 1998). Propionyl CoA carboxylase is involved in the catabolism of several branched chain amino acids and fatty acids of odd-carbon chain lengths by converting propionyl CoA to methylmalonyl CoA which ultimately enters the Krebs cycle. β-Methylcrotonyl CoA carboxylase participates in the catabolism of leucine. Disorders of each of the three mitochondrial enzymes give rise to specific carboxylase deficiencies (Wolf, 1995). A more severe disease state arises when the activity of all four biotin enzymes is compromised. This state is known as multiple carboxylase deficiency (MCD).

1.5.4.1 Multiple Carboxylase Deficiency

MCD is an inherited metabolic disorder which can take on one of two forms. Early onset or neonatal MCD is also referred to as holocarboxylase synthetase deficiency syndrome as this is a biotin ligase disorder. Here the biotinylation and hence activity of all biotin enzymes is compromised in patients at a very early age. An accumulation of the toxic metabolites arising from deficiencies of each of the four biotin carboxylases results in, amongst other symptoms, developmental delay (Wolf, 1995). The late onset or juvenile form of MCD arises due a deficiency in the supply of biotin to cells. Defects in biotinidase, the enzyme responsible for recycling protein-bound biotin as well as functioning as a serum biotin binding protein (Hymes & Wolf, 1999), have been identified as the cause of this form of MCD (Wolf, 1995). The age of onset of symptoms ranges from several weeks to several years of age (Wolf, 1995) and can also be induced through a low biotin diet and an excessive dietary intake of

the egg-white protein avidin in its non-denatured state (Parsons *et al.*, 1937). Both forms of MCD are, in most cases, treatable by oral supplements of 10-20 mg of biotin per day (Wolf, 1995).

Early onset MCD is an autosomal recessive disease. Recently the human *bpl* gene was mapped to a 2.1 Mb region of chromosome 21 which is also involved in the pathogenesis of Down's syndrome (Dahmane *et al.*, 1998). However it seems unlikely that BPL is involved in that disease as a number of developmental genes are also present in this region of the chromosome (Dahmane *et al.*, 1998).

To date, 10 point mutations in BPL which give rise to early onset MCD have been identified (Suzuki *et al.*, 1994; Aoki *et al.*, 1995; Dupuis *et al.*, 1996; Sakamoto *et al.*, 1998). Some of these are frameshift and missense mutations resulting in truncated, non-functional proteins. Most of the reported lesions in the *bpl* gene cluster to the presumptive biotin binding region based on homology with BirA (Figure 1.7) (Wilson *et al.*, 1992; Dupuis *et al.*, 1996; Chapman-Smith & Cronan, 1999a). Analysis of the activity of mutant BPLs in fibroblast extracts (Wolf, 1995; Burri *et al.*, 1981; Aoki *et al.*, 1997; Aoki *et al.*, 1999) have revealed that in all but one case, the lesions cause an increase in the K_m for biotin. The decrease in affinity for biotin is correlated with the age of onset of the disease with the more severe mutants inducing an earlier onset of symptoms (Burri *et al.*, 1981; Suzuki *et al.*, 1994; Aoki *et al.*, 1997; Aoki *et al.*, 1999). Interestingly the one mutation which does not effect the affinity for biotin, L237P, is one of two identified mutations which have been mapped to the N-terminal region of the enzyme outside the proposed catalytic domain (Figure 1.8) (Aoki *et al.*, 1997). Expression of a cDNA encoding this mutant BPL in fibroblasts and analysis of cellular extracts revealed that the V_{max} of the enzyme was decreased relative to

wildtype BPL but the K_m for biotin was unaffected (Aoki *et al.*, 1997). Further work is required to determine the actual mechanism by which this mutation affects BPL activity.

1.5.5 Yeast BPL

A good model system for understanding the structure, function and disease associated with human BPL appears to be the yeast *S. cerevisiae*. Firstly both organisms are incapable of synthesising biotin *de novo* and therefore do not contain the biotin synthetic regulation system observed in *E. coli*. Secondly the BPLs from human and yeast are similar in size (Figure 1.9) and share some sequence homology in their N-terminal regions which is not present in bacterial or plant BPLs (Figure 1.8).

Mishina and coworkers (1980) first proposed that only one BPL was present in *S. cerevisiae* and that this enzyme catalysed the biotinylation of all cellular enzymes. Based on a screen of over 7000 mutants requiring saturated fatty acids, two complementation groups were obtained (Mishina *et al.*, 1980). Members in the first complementation group, *acc1*, displayed the biotin carboxylase and transcarboxylase activities associated with ACC. One of these mutants was purified and confirmed to be ACC (Roggenkamp *et al.*, 1980). The second complementation group, *acc2*, contained an enzyme capable of activating both ACC and PC (Mishina *et al.*, 1980). The gene for the yeast BPL was isolated by complementation of an *E. coli birA*$^-$ strain and shown to be allelic to *acc2* (Cronan & Wallace, 1995). This gene was termed biotin protein ligase 1, *bpl1*. The presence of one *bpl* locus was first suggested by Sundaram and coworkers (1971) who purified a single species of BPL from yeast extract. BPL is

essential to yeast as disruption of one of the two *bpl1* alleles in diploid cells resulted in only half of the haploid spores surviving after sporulation (Hoja *et al.*, 1998). Of those spores which survived none were shown to be *bpl1*$^-$.

1.6 BIOTIN METABOLISM IN YEAST

Biotin metabolism in yeast and other eukaryotes is relatively poorly understood. The genes and proteins involved in the import, synthesis, and utilisation of biotin are beginning to be determined but the characterisation of these molecules and the processes they control are presently unclear.

The yeast *S. cerevisiae* is a biotin auxotroph which obtains biotin from two sources. Firstly exogenous biotin can be imported into the cell through the action of the recently cloned VHT-1 biotin/H^+ symporter (Stolz *et al.*, 1999). The activity of VHT-1 and the level of its mRNA are modulated by the concentration of extracellular biotin (Rogers & Lichstein, 1969b; Stolz *et al.*, 1999). Low concentrations of biotin in the media (0.8 nM) increase the level of VHT-1 mRNA and thus activity of the protein whilst higher biotin concentrations (8 nM) are inhibitory (Stolz *et al.*, 1999). The mechanism by which biotin regulates this transcription process is unclear. The second exogenous source of biotin is through the biosynthesis of the vitamin from the precursors 7-keto,8-amino-pelargonic acid (KAPA), 7,8-diamino-pelargonic acid and dethiobiotin. The genes for the enzymes responsible for converting these precursors into biotin have recently been cloned (Zhang *et al.*, 1994; Phalip *et al.*, 1999) along with the KAPA import protein (Phalip, *et al.*, 1999). All of these genes, apart from the biotin synthase gene, are present in an organised gene cluster (Phalip *et al.*, 1999). Regulation of gene expression from this operon remains to be determined.

Unlike mammalian organisms, yeast do not appear to recycle protein-bound biotin. The existence of a yeast homologue of mammalian biotinidase has not been reported. A recent search of the *S. cerevisiae* genome database failed to detect a coding region sharing homology with the cDNA for human biotinidase. Whether or not yeast contains an alternate mechanism for reusing protein-bound biotin is not known.

1.6.1 Yeast biotin enzymes

Yeast BPL (yBPL) is responsible for the *in vivo* biotinylation of five enzymes. These enzymes are ACC, the two isoforms of PC, PC1 and PC2, urea amidolyase and an as yet unidentified protein. ACC is an essential enzyme as disruption of the *acc1* gene is lethal to spores, even when grown with fatty acid supplementation (Hablacher *et al.*, 1993). This is an interesting observation as there has only been one report in the literature of a patient diagnosed with acetyl-CoA carboxylase deficiency (Blom & Scholte, 1981). Additionally, both human and yeast ACC both reside in the cytoplasm along with BPL which would provide every opportunity to ensure biotinylation of this essential enzyme.

Two groups have demonstrated that PC is also an essential enzyme but only when the activity of both isozymes is abolished (Stucka *et al.*, 1991; Walker *et al.*, 1991; Brewster *et al.*, 1994). Single knockout strains are viable but the contribution of each isoenzyme to total PC activity varies depending on the yeast strain used. Walker and coworkers demonstrated that PC1 contributes 80% of cellular PC activity in the DBY746 strain (Walker *et al.*, 1991; Brewster *et al.*, 1999). This is in contrast to the report of Stucka and coworkers (1991) who showed that PC1 and PC2 contribute equally to enzyme activity in the W303-1A

strain. It appears that PC1 is most responsible for supplying the TCA cycle with C4 intermediates during fermentative growth whilst the repression of PC2 in later stages of growth suggests a role for this isoenzyme during the establishment phase of growth (Brewster *et al.*, 1994).

The fourth known biotin enzyme in yeast is urea amidoylase. This enzyme consists of a single polypeptide chain and catalyses two reactions in the degradation pathway of urea. Firstly the carboxylation of urea to allophanate occurs through the attached biotin moiety. This is followed by the hydrolysis of allophanate to ammonia and carbon dioxide (Whitney & Cooper, 1972). The expression of urea amidolyase is induced by the presence of allophanate and repressed by nitrogen catabolites (Genbauffe & Cooper, 1986). Sequencing of the gene (Genbauffe & Cooper, 1991) has revealed the amino acid sequence around the biotin attachment site is not the classic Met-Lys-Met motif, but Met-Lys-Ala. Sumrada and coworkers (1982) noted that upon induction of the single chain polypeptide a time lag of 2 to 6 mins could be measured between the appearance of allophanate hydrolyase and urea carboxylase activities of the enzyme. This lag could be accounted for by the time required for biotinylation of the latter component of the enzyme.

1.7 PYRUVATE CARBOXYLASE

In the present study, the interaction between BPL and the biotin domain of PC1 from *S. cerevisiae* was investigated. A discussion on PC and its structure is contained in the following section.

PC catalyses the ATP dependent formation of oxaloacetate from pyruvate and HCO_3^-, as shown in Eq. 4.

$$\text{Enz-biotin} + \text{MgATP} + HCO_3^- + \text{pyruvate} \rightleftharpoons \text{Enz-biotin} + \text{MgADP} + \text{Pi} + \text{oxaloacetate} \qquad \text{(Eq. 4)}$$

The anaplerotic role of PC is to replenish oxaloacetate which is an intermediate in the Krebs cycle (Jitrapakdee & Wallace, 1999; Wallace *et al.*, 1998). The metabolites generated from this process participate in gluconeogenesis, lipogenesis, the synthesis of certain amino acids and the formation of neurotransmitters (Wallace *et al.*, 1998). Unlike mammalian PC, the yeast enzyme is cytosolic (Rhode *et al.*, 1991) and therefore co-localises with BPL. The carboxylation reaction catalysed by PC is dependent upon the specific attachment of a biotin moiety to the enzyme.

1.7.1 Primary and domain structure of PC

The first protein sequence obtained for PC came from a study on the biotin accepting site of avian and sheep enzymes (Rylatt *et al.*, 1977). Biotin was shown to be covalently attached to the ε-amino group of the side chain of a specific lysine residue present in an Ala-Met-Lys-Met motif. Around the same time, identical biotinylation motifs were also identified for *P. shermanii* TC and *E. coli* ACC (Sutton *et al.*, 1977; Wood & Barden, 1977). Deduced primary structure from the sequencing of partial cDNA clones of PC and other biotin enzymes revealed that the homology around the biotin attachment site extended beyond the tetrapeptide (Freytag & Collier, 1984; Lamhonwah *et al.*, 1986). Together these data suggested a common evolutionary origin for biotin accepting domain of biotin enzymes (Lynen, 1975; Samols *et al.*, 1988).

The first complete sequence of a biotin carboxylase gene to be reported was yeast PC1 (yPC1) (Lim *et al.*, 1988). The deduced translation product was 1178 amino acids having a molecular mass of 130.1 kDa. Specific sequences within yPC were shown to have significant homology to certain regions in other biotin carboxylases, lipoamide transferases and carbamyl phosphate synthetases. Combined with partial proteolysis of yPC, these data identified the presence of three functional domains present within the single polypeptide chain; an N-terminal ATP/HCO_3^- binding domain, a central pyruvate binding domain and a C-terminal biotin accepting domain (Figure 1.10; Lim *et al.*, 1988). Since this report, cDNA sequences for PC from a wide variety of species have been reported (reviewed Jitrapakdee & Wallace, 1999) and all contain the same three functional domains as yPC.

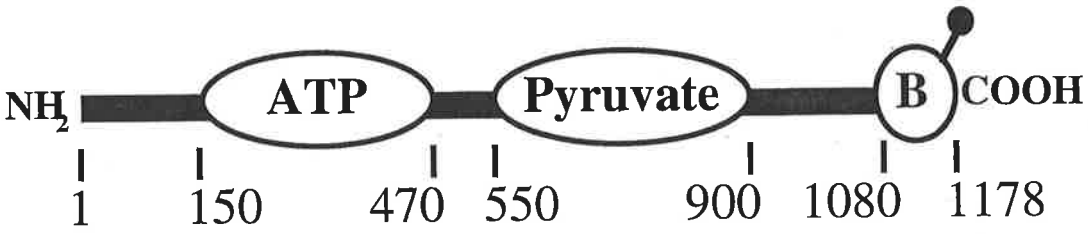


Figure 1.10: Domain structure of yPC. The relative positions of the ATP/HCO_3^- binding, pyruvate binding and biotin (B) accepting domains within yeast PC1 are shown (Lim *et al.*, 1988). The numbers given below represent the position of amino acids relative to the amino terminus. The biotin moiety attached in the C-terminal region is represented by the solid circle.

1.7.2 Quaternary structure of PC

The enzymatically active form of PC is a multimeric protein complex. The enzyme consists of four identical subunits of approximately 120-130 kDa (Wallace & Easterbrook-Smith, 1985) with the exception of PC from

Pseudomonas citronellolis (Goss *et al.*, 1981), *Azobacter vinelandii* (Scrutton & Taylor, 1974) and *Methanobacterium thermoautotrophicum* which are composed of two polypeptide chains of 75 kDa (α) and 52 kDa (β) (Mukhopadhyay *et al.*, 1998) arranged as an $(\alpha\beta)_4$ structure. The multimeric composition of PC is required for enzyme activity. Dilution studies on sheep (Khew-Goodall *et al.*, 1991) and chicken PC (Attwood *et al.*, 1993) revealed that loss of PC activity followed dissociation of the active tetrameric molecule into inactive dimers and monomers. The presence of acetyl CoA during the dilution of sheep PC preserved the tetrameric structure and consequently enzyme activity (Khew-Goodall *et al.*, 1991). Similarly, cold inactivation of chicken PC has been shown to occur through the dissociation of the active tetramer (Irias *et al.*, 1969). Sedimentation studies on yeast PC also showed that activity was only associated with a tetrameric enzyme (Taylor *et al.*, 1972; Taylor *et al.*, 1978). While it is clear from the above studies that the active form of PC is tetrameric, the reassociation of monomers into inactive tetramers suggests that the activity of PC is also dependent upon the correct conformation of the component monomers (Khew-Goodall *et al.*, 1991; Attwood *et al.*, 1993).

To date no high resolution structure of PC has been determined, mainly due to the difficulty researchers have encounted in growing crystals of the protein necessary for X-ray crystallography. Insights into the structure of PC have been derived from electron microscopy studies (Cohen *et al.*, 1979; Goss *et al.*, 1979; Mayer *et al.*, 1980; Johannssen *et al.*, 1983; Osmani *et al.*, 1984; Rhode *et al.*, 1986; Rhode *et al.*, 1991). Mayer and coworkers first proposed a definitive structure for chicken, sheep and rat PC (Mayer *et al.*, 1980). In this model, the enzyme conformation is that of a tetrahedron-like structure composed of two pairs

of subunits in two different planes orthogonal to each other with the opposing pairs of subunits interacting with each other on their convex surfaces (Figure 1.11). A cleft was observed running along the midline of the long axis. The biotin group, detected by analysing enzyme-avidin complexes, was located in this midline cleft on the concave monomer surface within 3 nm of the intersubunit junction (Johannssen *et al.*, 1983). In the presence of acetyl CoA this midline cleft becomes less visible and it has been proposed that this cleft area maybe the active site of enzyme (Mayer *et al.*, 1980). Subsequent investigations into other PC species has revealed that this structure is essentially common for all PC species (Osmani *et al.*, 1984; Rhode *et al.*, 1986) including the $(\alpha\beta)_4$ conformation of *P. citronellolis* PC (Fuchs *et al.*, 1988).

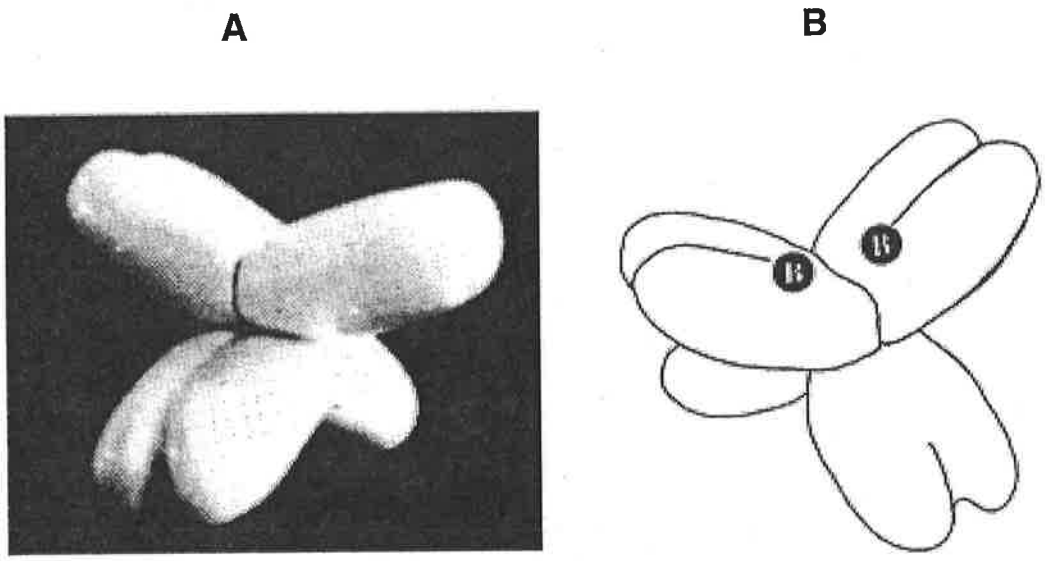


Figure 1.11: Structure of PC derived from electron microscope studies. A. PC consists of four subunits arranged as a tetrahedron structure with a midline cleft separating two distinct domains and running along the longitudinal axis of each monomer (reproduced from Mayer *et al.*, 1980). B. The position of the biotin moiety (solid circle) in the PC structure was determined by avidin binding studies (Johannssen *et al.*, 1983). Figure adapted from Val, 1995.

1.7.3 yPC Biotin Domains

The location of the biotin moiety along the midline cleft of the quaternary PC structure (Johannssen *et al.*, 1983; Fuchs *et al.*, 1988) implies that four biotin domains, which contain biotin and interact with BPL, reside in solvent exposed regions of the molecule. The primary sequence of the biotin domains of yPC1 and yPC2 have been reported (Lim *et al.*, 1988; Val *et al.*, 1995). Despite the high degree of amino acid homology within biotin domains, yPC1 and 2 appear to be unusual amongst the biotin carboxylases in that the biocytin residue is positioned greater than 35 residues away from the carboxyl terminus of the folded domain (Samols *et al.*, 1988; Chapman-Smith & Cronan, 1999a). This is due to an additional nine and ten C-terminal residues on yeast PC 1 and 2 respectively (Lim *et al.*, 1988; Val *et al.*, 1995). Val and coworkers showed that both biotin domains could be expressed as stable proteins in *E. coli* and folded into a conformation which was recognised and biotinylated by BirA (Val *et al.*, 1995). Deletion of the five C-terminal amino acids of yPC2 was shown to decrease the efficiency of *in vivo* biotinylation by up to 4-fold over a peptide containing this sequence (Val *et al.*, 1995). Similarly, the presence of additional primary structure in the N-terminal region of the yPC1 biotin domain also improved both expression and biotinylation (Val *et al.*, 1995). A peptide corresponding to the C-terminal 104 amino acids yPC1 was a 6-fold better substrate in *in vivo* biotinylation assays than a peptide encompasing only the C-terminal 85 amino acids (Val *et al.*, 1995). NMR experiments performed on the 85 and 104 amino acid residue biotin domain constructs from yPC1 showed that only the larger molecule was structured in solution (Chapman-Smith, pers.

comm.) suggesting that the 19 amino acids between 1075 and 1094 contribute to stabilising the domain structure

It is clear that *in vivo* biotinylation is an essential process which must occur in a reliable and specific manner in all organisms. The chemical attachment of biotin onto protein substrates, the availability of various techniques to detect this prosthetic group and use of yeast as a model for understanding biotin metabolism in humans makes investigation of this system a favourable one. Furthermore, the implication of BPL deficiency in patients with MCD imply that studies into the protein:protein interaction between this enzyme and its substrates is beneficial research.

1.8 PHAGE DISPLAY

1.8.1 Phage Display for Investigating Protein-Protein Interactions

Smith first demonstrated that a foreign protein could be displayed on the surface of an *E. coli* filamentous phage (Smith, 1985). This epitope was accessible to an antibody and the phage displaying the peptide could be selected from ordinary phage by immunoaffinity purification. Since those early experiments, phage display has developed into a powerful technique for investigating the interaction of proteins and peptides with other macromolecules (Clackson & Wells, 1994; O'Neil & Hoess, 1995; Dunn, 1996; Lowman, 1997). Polypeptides are expressed on the surface of filamentous bacteriophage as fusions to one of the coat proteins of the virion (Smith, 1985; Cesareni, 1992). The list of proteins that have been displayed on bacteriophage is ever increasing (Table 1.2) but include hormones such as growth hormone (Bass *et al.*, 1990), antibodies (reviewed Hoogenboom *et al.*, 1991; Winter *et al.*, 1994; Hudson, 1998; Vaughan *et al.*, 1998), proteases such as trypsin (Corey *et al.*, 1993) and variants of subtilisin (Atwell & Wells, 1999; Demartis *et al.*, 1999), nucleic acid binding proteins (reviewed O'Neil & Hoess, 1995; Friesen & Darby, 1997) and enzymes, such as β-lactamase (Soumillion *et al.*, 1994) and glutathione-S-transferase (Demartis *et al.*, 1999). Using phage display technology, proteins can be remodelled through mutagenesis and mutants with desirable phenotypes selected using an immobilised ligand. An understanding of the residues required for the function of the displayed protein can then be obtained by determining the primary structure of the selected mutants.

Table 2: Proteins Displayed on Bacteriophage

Protein	Fusion	Valency	Reference
Growth factors			
hGH	gIIIp	Mono	Bass *et al.*, 1990
IL3	gIIIp	Mono	Gram *et al.*, 1993
IL2	gIIIp	Mono	Buchli *et al.*, 1997
IGF binding protein 2	gIIIp	Mono	Lucic *et al.*, 1998
Antibodies			
Epitope	gIIIp	Poly	Smith, 1985
Single chain Fv	gIIIp	Mono	McCafferty *et al.*, 1990
Fab fragments	gIIIp	Mono	Hoogenboom *et al.*, 1991
Proteases			
Trypsin	gIIIp / gVIIIp	Poly	Corey *et al.*, 1993
H64A subtilisin	gIIIp	Mono	Demartis *et al.*, 1999
Protease Sensitive Substrates			
Factor Xa	gIIIp	Mono	Matthews & Wells, 1993
Furin	gIIIp	Mono	Matthews *et al.*, 1994
α-lytic protease	gIIIp	Mono	Lien, 2000
Protease Inhibitors			
BPTI	gIIIp	Poly	Roberts *et al.*, 1992
Chymotrypsin Inh.	gIIIp	Poly	Holmes *et al.*, 1996
Tendamistat	gIIIp	Mono	McConnell & Hoess, 1995
Nucleic acid Binding Proteins			
Zinc finger protein	gIIIp	Mono	Rebar & Pabo, 1994
Enzymes			
β-lactamase	gIIIp	Poly	Soumillion *et al.*, 1994
Alkaline phosphatase	gIIIp / gVIIIp	Poly	McCafferty *et al.*, 1991
Glutathione-S-transferase	gIIIp	Mono	Demartis *et al.*, 1999
Subtiligase	gIIIp	Mono	Atwell & Wells, 1999
Peptides			
Antibody binding	gIIIp	Poly	Cwirla *et al.*, 1990
Streptavidin binding	gIIIp	Poly	Caparon *et al.*, 1995
Biotin binders	gVIIIp	Poly	Saggio & Laufer, 1993

1.8.2 Bacteriophage structure

Filamentous bacteriophage comprise a family of 10 viruses which contain about 10 genes and infect Gram-negative bacteria. The wildtype virion is a flexible rod about 6 nm in diameter and, depending on the strain, 800 - 2000 nm long (reviewed in Cesareni, 1992; Makowski, 1994; Marvin, 1998). The rod functions as a sheath which surrounds a single stranded DNA molecule at the core. This sheath is composed of several thousand identical copies of an α-helical protein, the gene VIII product (gVIIIp) which are held together by hydrophobic interactions through a 19 residue apolar domain. The virion structure is capped at each end with a few copies of minor proteins. On the distal end of the virus, which is assembled first, there are five copies of each of the small hydrophobic peptides gVIIp and gIXp. At the proximal end of the virus are five copies of the gVIp and gIIIp proteins. The gIIIp is the protein responsible for binding to the F-pilus structure of the host cell in the first step of phage infection.

The gIIIp protein, as shown in Figure 1.12, has a modular structure consisting of three domains of 67 (N1), 130 (N2) and 123 (CT) amino acids connected by glycine rich linkers of 19 (G1) and 39 (G2) amino acids (Lubkowski *et al.*, 1998; Marvin, 1998). Under the electron microscope the N1-N2 structure appears as a knob at one end of the virus. Digestion of the virion with subtilisin cleaves gIIIp within G2, removing the knob structure and leaving the CT domain buried in the gVIIIp array (Marvin, 1998). Recently the structure of the N1 domain was determined by NMR (Holliger & Riechmann, 1997) and the structure of N1-N2 fusion determined by X-ray crystallography (Lubkowski *et al.*, 1998). These structures have provided insights into the mechanism of bacteriophage infection. The N2 domain absorbs to the tip of the bacterial conjugative pilus that retracts

by an unknown process and this draws the virion into contact with the cell (Marvin, 1998). At this stage the interaction of the gIIIp N1 domain with the bacterial outer-membrane protein TolA (Riechmann & Holliger, 1997) permits the entry of the phage DNA into the cytoplasm by an unknown mechanism.

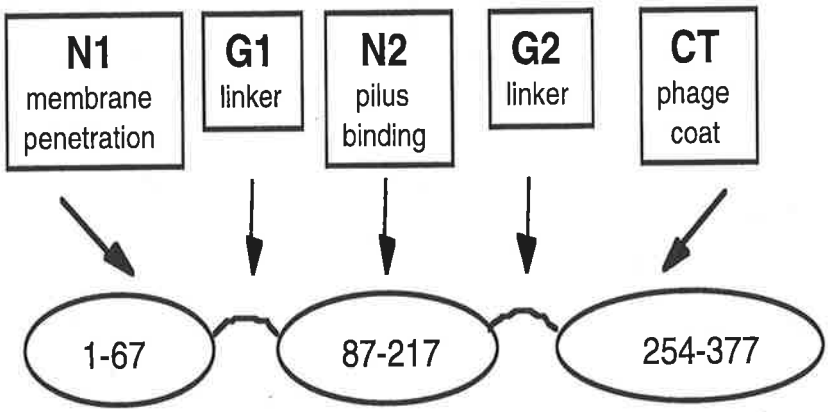


Figure 1.12: Domain structure of gIIIp. The three domains of gIIIp (N1, N2 and CT represented by circles) are connected by glycine rich linkers (single lines). The amino acid residues found in each domain are shown. The functions associated with each of the regions of gIIIp is shown above the protein.

1.8.3 Fusion of Proteins to gIIIp

The display of foreign proteins on the surface of filamentous bacteriophage requires the fusion of the target molecule to a virion coat protein. Two phage proteins have been manipulated for this process; the major coat protein (gVIIIp) and the minor coat protein (gIIIp) (Table 1.2). Due to the specific packaging requirements of the gVIIIp molecule for the assembly of bacteriophage, its use appears generally to be limited to the display of small peptides on the surface of virions (Greenwood *et al.*, 1991; Makowski, 1994) although the display of antibodies has been reported (Hoogenboom *et al.*, 1991; Kang *et al.*, 1991). Additionally, the introduction of a single copy of the foreign protein on the virus, or monovalent display, is technically difficult. Monovalent

display is often desirable when engineering proteins with altered binding properties as it avoids the problems of avidity during the selection process.

The gIIIp is an ideal vehicle for the monovalent display of foreign proteins. Monovalent phage display avoids the problem of reduced infection often observed with multivalent display where each copy of the gIIIp molecule is fused with a foreign protein (Bass *et al.*, 1990; Lowman & Wells, 1991). Here the DNA encoding the foreign protein is fused to the gene 3 and is expressed under the control of a weak promoter from a plasmid that is independent of the phage. This vector contains the origins of replication for both *E. coli* and M13 phage. By infecting bacterial cells containing the construct with helper phage infectious phagemid particles are produced, displaying on average less than one copy of the foreign protein and harbouring single-stranded phagemid DNA.

The power of phage display lies in the fact that there is a covalent link between the mutant protein and the DNA which encodes it. Remodeling of proteins begins with mutagenesis on the DNA encoding the displayed protein. Both the mutant DNA, which can then be sequenced to determine the nature of the mutation, and the mutant protein, can be isolated together as a phagemid particle. The selection of phage displaying mutants with desirable characteristics is achieved by biopanning. Here an interacting capture molecule is immobilised to a surface and those phage interacting with the capture molecule are collected while the non binders are washed away. Mutants with desirable characteristics are enriched through repeated cycles of infection, production of new phage and biopanning. Phage display is an attractive tool for studying protein:protein interactions as large libraries of mutant proteins can be generated and screened in a rapid and relatively inexpensive manner.

1.9 PROJECT RATIONALE

An objective of our laboratory over the years has been to investigate the protein structure and function relationships of PC. Obviously the post-translational biotinylation of PC is a key event in the expression of an active gene product. Biotinylation *in vivo* is a ubiquitous process, essential for the maturation of certain enzymes which catalyse important metabolic reactions. Therefore, biotinylation must occur reliably and specifically in all living cells. We were interested in understanding this process further using yeast BPL and the biotin domain of PC1 as a model system.

At the time this project commenced the *bpl* gene from *S. cerevisiae* had been isolated and sequenced in our laboratory by Prof. John Cronan (Cronan & Wallace, 1995). The encoded protein was similar in size to the product of the human gene (Leon-Del-Rio *et al.*, 1995; Suzuki *et al.*, 1994) but was twice the size of the BirA (Howard *et al.*, 1985). Sequence homology in the C-terminal halves of the family members was observed but the structure and function of the N-terminal region of the eukaryotic enzymes was unclear. Recombinant expression and purification of a eukaryotic BPL, necessary for further investigation of the enzyme, had not been reported. In addition, a three dimensional structure of the protein substrate for BirA had not been reported. Thus interpretation of mutational studies performed upon both BPLs and various biotin domains was difficult and ambiguous.

An ultimate goal of this study is to understand the macromolecular interaction which occurs between BPLs and biotin domains.

The specific aims of this thesis were to:

1) Recombinantly express yeast BPL in *E. coli* and purify enzymatically active protein.

2) Determine the assay conditions required to perform *in vitro* biotinylation necessary for the kinetic analysis of substrate binding.

3) Using limited proteolysis, determine the domain structure of the enzyme especially in the N-terminal region of the enzyme. As the presumed catalytic region of yeast BPL maps to the C-terminal region, a series of truncated forms of BPL can be generated to assess the importance of the N-terminal region in catalysis.

4) Produce phagemid particles displaying a large number of randomly mutated yeast PC1 biotin domains as a fusion to the gIIIp coat protein. Screen the population of phagemids to select those mutants which interact with BPL and are subsequently biotinylated.

5) Segregate the mutants selected by phage display into those biotin domains which have a lowered affinity for BPL using an *in vivo* screen in *E. coli*.

6) Purify suitable biotin domain variants for kinetic analysis with BirA and the yeast enzyme purified and characterised above.

Chapter 2

MATERIALS AND METHODS

MATERIALS AND METHODS

2.1 CHEMICALS

The following chemicals were obtained from Sigma Chemical Co., St Louis, MO., USA: agarose (type 1), ampicillin, ATP (disodium, grade I), β-mercaptoethanol, bovine serum albumin (BSA), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (BCIG), 5-bromo-4-chloro-3-indolyl-phosphate (BCIP), d-biotin, Coomassie Brilliant Blue (G_{250} and R_{250}), dithiothreitol (DTT), ethidium bromide, ethylenediaminetetraacetic acid (EDTA), 3-[N-Morpholino]propane-sulfonic acid (MOPS), nitroblue tetrazolium, phenylmethylsulphonyl fluoride (PMSF), polyoxyethylene-sorbitan monolaurate (Tween 20), sodium dodecyl sulfate (SDS), glycyl glycine, N,N,N',N'-tetramethylethylenediamene (TEMED), Tris-(hydroxymethyl)-methylamine (Tris) and Triton X-100. Acrylamide and bisacrylamide were purchased from BioRad Laboratories Inc., Herates, CA, U.S.A. Superose 6 and Sepharose CL-6B were purchased from Pharmacia, Uppsala, Sweden. Other fine chemicals were purchased from Merck Pty. Ltd., Kilsyth, Vic., Australia.

2.2 RADIOCHEMICALS

[α-^{32}P] dATP and [γ-^{32}P] ATP were purchased from Bresatec Ltd., Adelaide, South Australia. *d*-[8,9-^{3}H]biotin was purchased from Amersham Australia, North Ryde, NSW, Australia.

2.3 OLIGONUCLEOTIDES

All oligonucleotides used in this study were purchased from Bresatec Ltd., Adelaide, South Australia. The restriction sites in the oligonucleotides are underlined and mutagenic changes are in bold. The sequences of the oligonucleotides are as follows :

YPC104B 5'- CATACCATGGCAATGAGAAAAATTCGTGTTGCTG-3'

YPC3'B 5'-TTACTGCAGACTATGCCTTAGTTTCAACAGGAACT-3'

Lys61Leu

5'-AAGATGTGATCATTTCCAT**GAG**CATGGCGCTTAATACGGCTA-3'

YPC104C 5'-ACAGAATTCCATGGCAATGAGAAAAATTCGTGTTGCTG-3'

YPC-3'C 5'-GCACAGCACCACCTGCAGACTA-3'

Bpl5' 5'-CAACTA**TC**ATGAATGTATTAGTCTATAATGGC-3'

Bpl3' 5'-CCATTGTAGGGTCACCTTGAGC-3'

BplPst 5'-TCAGAAAAGCTGCAGGCACTC-3'

BplBam 5'-AAAGTC**GGATCCTAATGATGATGATGATGATG**

ACTCTGAACCTTTTTAGCAATTAAGC-3'

BplE233 5'-ATCATACC**ATGG**AAACCGTTGTGGAAAACCTG-3'

BplE369 5'-ATCATACC**ATGG**AATACTTCAAGTATCTGAATGTTC-3'

BplE409 5'-ATATAACC**ATGG**AAAGCACTTTACTTCACGTGGG-3'

BplPstBack 5'-GAGTGCCTGCAGCTTTTCTGC-3'

AraFor 5'-TGCCTGACGGTTTTTGCC-3'

AraFor2 5'-AACTATGGCTGGAATGTCC-3'

forward sequencing primer 5'-GTAAAACGACGGCCAGT-3'

reverse sequencing primer 5'-CACACAGGAAACAGCTATGACCATC-3'

2.4 MOLECULAR BIOLOGY KITS

The *fmol*™ DNA sequencing system and TA cloning system were purchased from Promega, Madison, WI, USA. The BRESA-Spin™ Gel Extraction and Plamsid Mini kits were purchased from Bresatec Ltd., Adelaide, South Australia. The Geneclean II kit was obtained from BIO 101, CA, U.S.A..

2.5 ENZYMES AND PROTEIN MARKERS

All restriction enzymes were purchased from Pharmacia, Uppsala, Sweden and Bresatec Ltd., Adelaide, South Australia. T4 DNA ligase, T4 DNA polymerase, T4 polynucleotide kinase, and *Taq* DNA polymerase were purchased from Promega, Madison, WI, U.S.A. *Pwo* DNA polymerase was purchased from Boehringer Mannheim. RNase A and avidin-conjugated with alkaline phosphatase were purchased from Sigma. NiNTA-conjugated with alkaline phosphatase was purchased from Qiagen. Biotinylated protein markers were obtained from BioRad. Mark12™ wide range protein standards were purchased from Novex, San Diego, CA, USA. Full range Rainbow protein molecular weight markers were purchased from Amersham Australia, North Ryde, NSW, Australia. Prestained protein molecular weight markers were purchased from New England BioLabs, CA, USA.

2.6 BACTERIAL STRAINS:

E. coli DH5α: *sup*E44Δ*lac*U169(p80*lac*ZΔM15) *hsdR17* *rec*A1 *end*AA1 *gyr*A96 *thi*-1 *rel*A1 host for recombinant plasmids (New England, Biolabs, CA).

E. coli XL1-Blue: *recA1 end A1 gyrA96 thi-1 hsdR17 SupE44 relA1 lac* [F' *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^{r})] for phage display (Stratagene, La Jolla, CA, USA).

E. coli BL21: *E. coli* B F^{-} *dcm ompT hsdS*(r_B^- m_B^-) *gal* for expression of recombinant proteins (Stratagene, La Jolla, CA, USA).

E. coli BL21(λDE3): *E. coli* BL21 carrying (λDE3) insertion, for expression of recombinant proteins using pET expression vectors (Stratagene, La Jolla, CA, USA).

The *E. coli* strains used for complementation assays were the *birA1 bioC* strain CY918 (Chapman-Smith *et al.*, 1994) and the *birA85 bioC* strain BM4062 (Barker & Campbell, 1981b).

2.7 BACTERIAL GROWTH MEDIA

LB broth: 1% (w/v) Bacto-tryptone (Difco), 0.5% (w/v) yeast extract (Difco), 1% (w/v) NaCl, adjusted to pH 7.0 with NaOH.

LB agar plates were made by adding 1.5% (w/v) Bacto-agar (Difco) to the LB broth

2YT broth: 1.6% (w/v) Bacto-tryptone (Difco), 1% (w/v) yeast extract (Difco), 0.5% (w/v) NaCl, adjusted to pH 7.0 with NaOH

Rich broth: 1% (w/v) Bacto-tryptone (Difco), 0.05% (w/v) yeast extract (Difco), 1% (w/v) NaCl, adjusted to pH 7.0 with NaOH

RB agar plates were made by adding 1.2% (w/v) Bacto-agar to Rich broth

Long term storage of plasmids in bacteria were prepared by adding 40% glycerol to an overnight culture of *E. coli* in LB-broth and then stored at -80°C.

2.8 PLASMIDS

1. pBluescript II (KS$^-$); Stratagene, La Jolla, CA, USA.

2. pET-16b; Novagen, Madison, WI, USA.

3. pKK223-3; Pharmacia, Uppsala, Sweden

4. pGF-14, a derivative of pHEN-1 (Hoogenboom *et al.*, 1991), was provided by Dr Göran Forsberg, Department of Biochemistry, University of Adelaide, Adelaide.

2.9 DNA TECHNIQUES

2.9.1 Small scale preparation of plasmid DNA

A single colony of *E. coli* was grown in 3 ml of LB broth containing appropriate antibiotic(s) and incubated at 37°C with shaking overnight. The bacterial culture was centrifuged at 13,000 rpm for 2 min. The supernatant was discarded and the pellet was resuspended in 100 ml of TES solution [25 mM Tris-HCl, pH 8.0, 10 mM EDTA pH 8.0, 15% (w/v) sucrose] and left on ice for 5 min. Two hundred microlitres of lysis solution [1% (w/v) SDS in 0.2 M NaOH] were then added and left on ice for 5 min. The lysates were then neutralised with 150 ml of 3 M potassium acetate pH 5.2 and left on ice for 5 min. Cellular debris and chromosomal DNA were removed by centrifugation at 13,000 rpm for 15 min. The supernatant was transferred to a new Eppendorf tube and RNA removed by digesting with 40 µg of RNAse A at 37°C for 30 min before being extracted twice with equal volumes of phenol and chloroform [1:1 (v/v)] The aqueous phase was separated from the organic phase by centrifugation at 13,000 x g for 2 min. The aqueous phase was then transferred to a new tube and extracted once with an

equal volume of chloroform. The plasmid DNA was precipitated by adding 2 volumes of cold absolute ethanol and kept at -20°C for 1h. DNA was recovered by centrifugation at 13,000 x g for 15 min and washed once with 70% ethanol. The pellet was dried under vacuum and dissolved in 30 µl sterile distilled water.

2.9.2 Large scale preparation of plasmid DNA by phenol/chloroform/PEG precipitation

A single colony of *E. coli* was grown with vigorous shaking at 37°C for 6 h in 10 ml of LB broth containing 100 µg/ml ampicillin. Five millilitres of culture were then added to 250 ml of the same culture medium and grown at 37°C with vigorous shaking at 37°C overnight. The cells were harvested by centrifugation at 5000 x g at 4°C for 10 min. Cells were suspended in 6 ml of TES buffer supplemented with 2 mg/ml lysozyme and left on ice for 30 min. Freshly prepared 0.2 M NaOH-1% (w/v) SDS (12 ml) was added, gently mixed and left on ice for 10 min. The lysates were neutralised by adding 7.5 ml of 3 M potassium acetate pH 4.8, gently mixed and incubated on ice for 10 min. Cellular debris and chromosomal DNA were removed by centrifugation at 10,000 x g for 15 min at 4°C. The supernatant was extracted twice with an equal volume of phenol/chloroform [1:1 (v/v)], and once with an equal volume of chloroform. The supernatant, containing plasmid DNA was precipitated by adding 2 volumes of cold absolute ethanol and kept at -20°C for 1 h. The plasmid was recovered by centrifugation at 10,000 x g for 15 min at 4°C. The pellet was washed once with 70% ethanol and dried under vacuum before being dissolved in 1.6 ml sterile distilled water. The plasmids were further purified by adding 400 µl of 4 M NaCl

and 800 µl of 33.3% polyethylene glycol (PEG) 6000, left on ice for 1 h and centrifuged at 10,000 x g for 15 min. The pellet was washed once with 70% ethanol and dried under vacuum. The dried pellet was then dissolved in 400 ml of sterile distilled water and extracted twice with an equal volume of phenol/chloroform [1:1 (v/v)] and once with an equal volume of chloroform. The plasmid was precipitated by adding 2 volumes of cold absolute ethanol and kept at -20°C for 1 h and centrifuged at 10,000 x g for 15 min. The pellet was then washed once with 70% ethanol, dried and dissolved in 0.5 ml sterile distilled water.

2.9.3 Purification of DNA fragment from agarose gel using the Geneclean kit

A slice of agarose containing a DNA fragment was excised under UV-transillumination and put in a 1.5 ml microtube. A half volume of TBE modifier and 4.5 volumes of NaI solution were added to a slice of gel and incubated at 65°C for 5 min. Glassmilk, 5 µl, was added and the DNA allowed to bind at 4°C for 5 min. Glassmilk-bound DNA was pelleted by centrifugation at 13,000 x g for 10 sec and washed 3 times with New Wash solution. DNA was eluted from Glassmilk with 20 µl of water by incubation at 65°C for 5 min. The Glassmilk was pelleted, as above, and the supernatant was kept.

2.9.4 Ligation

Plasmid DNA and the DNA insert were digested with restriction enzymes that generated compatible ends. The ligation was routinely carried out in a total

volume of 20 µl containing DNA insert : vector = 3 : 1 (molar ratio), 1X ligation buffer (66 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 1.0 mM DTT, 1.0 mM ATP) and 2 U of T4 DNA ligase at 14°C overnight.

2.9.5 Bacterial transformation

A single colony of *E. coli* was grown in 2 ml of LB broth at 37°C overnight. The overnight culture (600 µl) was then subcultured into 50 ml of LB broth and grown to A_{600} of 0.5-0.6. Cells were pelleted by centrifugation at 5000 x g for 10 min and suspended in 30 ml of 200 mM $CaCl_2$ and incubated on ice for 30 min. Cells were pelleted by centrifugation at 5000 x g for 10 min and resuspended in 3 ml of 80 mM $CaCl_2$ and incubated on ice for at least 60 min.

Typically, 200 µl of competent cells were mixed with 10 ml of ligation reaction and left on ice for 60 min. DNA and cell mixture were heat-shocked at 42°C for 3 min. The transformed cells were recovered by incubation in 1 ml of LB broth at 37°C for 45 min before plating. The recombinant clones were distinguished from non-recombinant clones on the basis of the α-complementation of β-galactosidase (lac Z) gene expression, using LB plates containing IPTG as the inducer and BCIG as substrate of β-galactosidase. Blue colonies indicate the non-recombinant clones while white colonies indicate the recombinant clones due to a disruption of the β-galactosidase gene on the plasmid.

2.9.6 DNA sequencing

2.12.6.1 fmol™ DNA sequencing

The 5'-end of the sequencing primer was labelled in a 10 ml reaction containing 100 ng of primer, 1x T4 polynucleotide kinase buffer (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 5.0 mM DTT, 0.1 mM spermidine) and 50 mCi of [γ-^{32}P]ATP. The reaction was carried out at 37°C for 30 min before inactivating the enzyme by heating at 95-100°C for 2 min. A master mix was then prepared including 1-2 µg of plasmid DNA, 5 µl of 5X sequencing buffer (250 mM Tris-HCl, pH 9.0, 10 mM $MgCl_2$), 1.5 µl of labelled primer and 5U of sequencing grade *Taq* polymerase. The reaction was brought up to a final volume of 17 µl with water. Master mix (4 µl) was transferred to 0.5 ml microtubes containing 2 µl each of ddATP, ddTTP, ddCTP and ddGTP separately and overlaid with 100 µl of mineral oil. The reaction was subjected to 30 cycles of amplification. Each cycle consisted of denaturation at 95°C for 30 sec (2 min for the first cycle), annealing at 50°C for 30 sec and extension at 70°C for 1 min. In cases where the length of primer was more than 24 nucleotides and / or with GC content greater than 50%, annealing and extension were performed at 70°C for 30 sec. Other parameters were essentially same as described above. The reaction was terminated by adding 4 µl of stop solution (10 mM NaOH, 95% formamide, 0.05% bromphenol blue, 0.05% xylene cyanole) and heated at 100°C before electrophoresis.

2.9.6.2 Dye Terminator sequencing

Plasmid DNA, isolated as described in 2.9.1, was further purified by adding 1 µl of 100 µg/ml RNase A to 20 µl of plasmid sample and incubated at

37°C for 30 min. Twenty microlitres of 7.5 M ammonium acetate and 40 µl of isopropanol were added, mixed and left at -20°C for 1 h. The tube was centrifuged at 13,000 x g for 15 min. The pellet was washed once with 300 µl of 75% ethanol in 50 mM sodium acetate and once with 100% ethanol, then dried and dissolved in 20 µl of water. The sequencing reaction was performed in a single 0.5 ml tube containing 1 µg of plasmid template, 8 µl of dye terminator mix and 100 ng of primer. The reaction was adjust to a final volume of 20 µl and overlaid with 100 µl of mineral oil before subjected to PCR. The PCR profile consisted of 25 cycles of denaturation at 96°C for 30 sec, annealing at 50°C for 15 sec and extension at 60°C for 4 min. After the thermocycling was complete, mineral oil was removed and sequencing products were precipitated by adding 0.1 volume of 3 M sodium acetate and 2 volumes of absolute ethanol, followed by centrifugation at 13,000 x g for 15 min. The pellet was dried under vacuum and submitted to the sequencing service at the Institute of Medical and Veterinary Science, Adelaide.

2.9.7 Quantitation of DNA by spectrophotometry

Plasmid DNA and single-stranded oligonucleotides were quantitated by measuring the absorbance at 260 nm. Samples were routinely measured in a 1 ml volume.

concentration of plasmid DNA (µg/ml) = A_{260} x 50 x dilution factor

concentration of oligonucleotides (µg/ml) = A_{260} x 20 x dilution factor

2. 10 PROTEIN TECHNIQUES

2.10.1 SDS-PAGE

Protein solutions containing 1-50 µg of total protein were suspended in an equal volume of 2X SDS Load buffer (0.125 M Tris-HCl, 4% (w/v) SDS, 10% (v/v) glycerol, 0.1% Bromophenol Blue) containing 5% β-mercaptoethanol and subjected to reducing discontinuous SDS-PAGE on a 4% stacking gel and 10-15% polyacrylamide separating gel. For optimal resolution of large molecular weight proteins glycine running buffer (0.125 M Tris-HCl, 1 M glycine, 0.1% SDS, pH 8.8) was used (Laemmli, 1970). For optimal resolution of small molecular weight proteins tricine running buffer (0.125 M Tris-HCl, 1 M tricine, 0.1% SDS, pH 8.3) was used (Schagger & von Jagow, 1987). Visualisation of protein was performed by soaking gels in Coomassie Blue Staining solution (10% (v/v) acetic acid, 0.3% (w/v) Coomassie Brilliant Blue R250) and destaining with 10% (v/v) acetic acid.

2.10.2 Western Transfer and Probing with Alkaline Phosphatase Conjugates

Following PAGE, the fractionated proteins were transferred to nitrocellulose filters using a semi-dry electroblotter (Multiphor II Novablot, Pharmacia, Uppsala, Sweden) with transfer buffer (20 mM Tris-HCl, 0.15 M Glycine, pH 7.5). For the detection of biotinylated proteins using avidin-alkaline phosphatase conjugate, the method described by Lim *et al* (1987) was employed. The detection of proteins containing a His6 tag used essentially the same protocol, except that PBS was replaced with Tris buffered saline (10 mM

Tris-HCl, 150 mM NaCl, pH 7.5) and the blot was probed using a 1/1000 dilution of NiNTA alkaline phosphatase conjugate (Qiagen).

2.10.3 Protein Quantitation

The concentration of protein solutions was determined by Bradford assays (Bradford, 1976). The concentration of solutions containing purified biotin domains was determined by microbiuret assay (Munkres & Richards, 1965). Quantitation of protein after PAGE was carried out by laser densitometry, using a Molecular Dynamics Model 300A densitometer with ImageQuant software (Sunnyvale, CA, USA), with adjustment for the effect of molecular weight on Coomassie Blue staining.

Electroblotting of peptides onto polyvinylidene difluoride was performed as described by Matsudaira (Matsudaira, 1988). N-terminal sequencing of proteins by automated Edman degradation was performed by Ms Dennise Miller using a Perkin-Elmer Procise 492 protein sequencer.

Chapter 3

EXPRESSION AND PURIFICATION OF YEAST BIOTIN PROTEIN LIGASE

EXPRESSION AND PURIFICATION OF YEAST BIOTIN PROTEIN LIGASE

3.1 INTRODUCTION

The gene encoding BPL from *Saccharomyces cerevisiae* (yBPL) was isolated and sequenced in our laboratory by Prof. John Cronan during his sabbatical. This gene was obtained by complementation using a conditionally lethal mutant strain of *E. coli* defective in biotin biosynthesis and BirA function (Cronan & Wallace, 1995). A cDNA for human BPL was also obtained by complementation (Leon-Del-Rio *et al.*, 1995) suggesting that bacteria posses the machinery to correctly fold at least small amounts of the relatively large eukaryotic BPLs. Sundaram and coworkers (1971) previously reported the purification of endogenous BPL from yeast but the multiple purification steps and low abundance of the enzyme combined to give low yields. Additionally, the protein obtained was not homogeneous. The emergence of recombinant DNA technology has permitted the overexpression of heterologous proteins in suitable hosts thus facilitating improved methods for the purification of proteins. This chapter describes the development of a purification strategy for yBPL recombinantly expressed in *E. coli* .

3.2 SPECIFIC METHODS

3.2.1 Expression of yeast BPL and preparation of cell lysates

Bacterial cultures of *E. coli* BL21(λDE3)pLys harbouring pET(Bpl-His_6) were grown in shake flasks in 2YT supplemented with 100 μg/ml ampicillin and 30 μg/ml chloramphenicol. Overnight cultures were diluted 1:100 into 2 l fresh media and grown at 30°C to A_{600} 0.6-0.8 before addition of IPTG to a final concentration of 0.1 mM. After 3 h, the cells were harvested by centrifugation, washed in binding buffer (20 mM Tris-HCl, pH 7.9, 0.5 M NaCl, 5 mM imidazole) and the pellet stored at -80°C overnight. The cell pellet was thawed on ice in 60 ml binding buffer, with the addition of 1 mM PMSF, resulting in cell lysis by the action of T4 lysozyme expressed from the co-resident pLysS plasmid. The disrupted cell suspension was sonicated, and centrifuged at 10,000 x g for 10 minutes. After a second centrifugation, the supernatant was filtered through a 0.45 μm Minisart filter (Sartorius) prior to chromatography.

3.2.2 Purification of yeast BPL

His6-tagged material was purified on a 2.5 ml Ni-NTA-agarose (Qiagen) gravity fed column. After loading the cell lysates onto the charged column equilibrated in binding buffer, the column was washed with 12 column volumes of binding buffer, 12 volumes of wash buffer (20 mM Tris-HCl, pH 7.9, 0.5 M NaCl, 10 mM imidazole) and bound material was eluted with 6 volumes of elution buffer (20 mM Tris-HCl, pH 7.9, 0.5 M NaCl, 0.5 M imidazole). Fractions containing yBPL were pooled and dialysed overnight against 4 l storage buffer

(50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol, 5% (v/v) glycerol). The BPL containing fractions eluted from the nickel column were further fractionated on a Q-Sepharose column (Pharmacia; 12 cm x 2.6 cm) with a 450 ml gradient from 0 to 250 mM NaCl in storage buffer, run at 5 ml/min. Fractions containing yBPL, detected by SDS-PAGE and Ni-NTA Western blot, were pooled and stored at -80°C.

3.2.3 Assay of yeast BPL

BPL activity was assayed by measuring the incorporation of [^{3}H]biotin into either apo BCCP-87 or apo yPC-104 as described by Chapman-Smith *et al* (1999). Briefly, the reactions contained 50 mM Tris-HCl, pH 8.0, 3 mM ATP, 5.5 mM $MgCl_2$, 5 μM biotin, 5 pmol [^{3}H]biotin (specific activity 35-44 Ci/mmol) 0.1 mM dithiothreitol, 0.1 mg/ml BSA and either 20 μM apo BCCP-87 or 5μM apo yPC-104. The reaction was initiated by addition of purified yBPL to a final concentration of 13 nM, except where stated otherwise, and incubated at 37°C for up to 30 minutes when aliquots of the reaction were spotted onto biotin- and trichloroacetic acid-treated filters. After air drying, the filters were washed twice in 10% ice cold trichloroacetic acid and once in ethanol, dried and the acid insoluble radioactivity measured. A unit of BPL activity was defined as the amount of enzyme required to incorporate 1 nmol of biotin per minute. Values for K_m and V_{max} were determined by fitting a plot of substrate concentration against rate to the Michaelis-Menten equation using GraphPad Prism for MacIntosh (GraphPad Software Inc, San Diego, CA). In some experiments, to obtain sufficiently high levels of radioactivity for accurate detection, it was necessary to

continue until greater than 10% of the limiting substrate had been utilised. In this case the data were transformed for altering substrate concentration by the method of Lee and Wilson (Lee & Wilson, 1971) and plotted as transformed values *s'* and *v'*.

3.3 RESULTS and DISCUSSION

3.3.1 Construction of yBPL and yBPL-His expression vector

The cloning strategy for the construction of a recombinant expression vector for the production of yBPL in *E. coli* is described here and is also outlined in Figure 3.1. Initially the DNA encoding yBPL was subcloned into pBluescript KS⁻II for subsequent cloning steps and as the template for PCR-based procedures. This plasmid, pKS(bpl), was produced by ligating a 2.6 kb *Bst*BI / *Sna*BI DNA fragment, containing the Bpl coding region, from pCY248 (Cronan & Wallace, 1995) into *Acc*I / *Sma*I digested pBluescript KS-II. The introduction of the *bpl* coding region into the pET16b bacterial expression vector (Novagen) required two cloning steps. Firstly, the 2.1 kb *Nco*I/*Bam*HI fragment excised from pKS(bpl), containing the *bpl* coding region from +131 bp to the end of the gene, was cloned into *Nco*I / *Bgl*II digested pET16b yielding the plasmid pN/B-Bpl. In order to fuse an initiation codon as well as the 5' +131 bp of coding sequence to the remaining 3' coding sequence, a suitable restriction site at the 5' end of the gene needed to be introduced. A *Bsp*HI site was engineered around the initiation codon by PCR mutagenesis using the primers Bpl5' and Bpl3' with pKS(bpl) as the template (Figure 3.2A). This product was ligated into pGEM-T (Promega) to produce pGEM(5'bpl). The 308 bp fragment liberated by digestion of this plasmid with *Bsp*HI and *Bst*EII was cloned into *Nco*I / *Bst*EII treated pN/B-Bpl. This yielded the expression vector pET(Bpl) where the *bpl* gene was under the control of the T7 promoter.

In order to facilitate downstream purification of the recombinant enzyme using nickel chelating chromatography, a vector was constructed for the

Figure 3.1: Cloning strategy for the construction of a bacterial expression vector for producing recombinant yBPL-His

The strategy for constructing pET(BPL-His) is described in 3.3.1 and outlined here. Abbreviations: ori, bacterial origin of replication; Amp R, ampicillin resistance gene; lacI, lactose operon repressor protein; MCS, multiple cloning site; pT7, T7 RNA polymerase binding site. The position of restriction sites employed in the cloning strategy are shown on the outside of the plasmids.

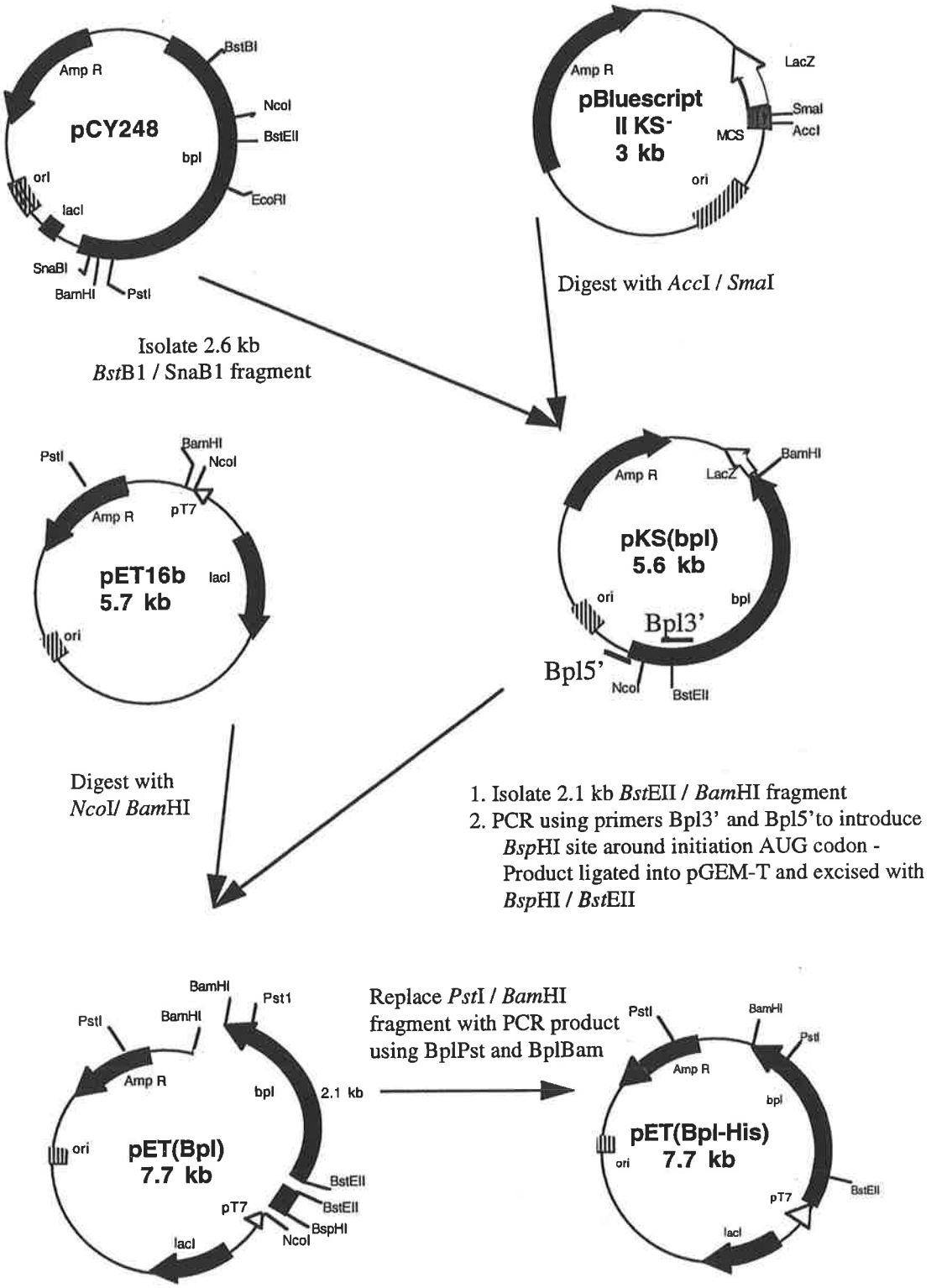

Amp R
BstBI
NcoI
BstEII
EcoRI
pCY248
bpl
ori
lacI
SnaBI
BamHI
PstI
Isolate 2.6 kb
*Bst*B1 / SnaB1 fragment
Amp R
LacZ
pBluescript
II KS-
3 kb
SmaI
AccI
MCS
ori
Digest with *Acc*I / *Sma*I
PstI
BamHI
NcoI
Amp R
pT7
pET16b
5.7 kb
lacI
ori
Amp R
LacZ
BamHI
pKS(bpl)
5.6 kb
ori
bpl
Bpl3'
Bpl5'
NcoI
BstEII
Digest with
*Nco*I/ *Bam*HI
1. Isolate 2.1 kb *Bst*EII / *Bam*HI fragment
2. PCR using primers Bpl3' and Bpl5'to introduce *Bsp*HI site around initiation AUG codon - Product ligated into pGEM-T and excised with *Bsp*HI / *Bst*EII
BamHI
Pst1
PstI
BamHI
Amp R
bpl
2.1 kb
ori
pET(Bpl)
7.7 kb
BstEII
BstEII
pT7
BspHI
NcoI
lacI
Replace *Pst*I / *Bam*HI
fragment with PCR product
using BplPst and BplBam
PstI
BamHI
Pst
Amp R
bpl
pET(Bpl-His)
7.7 kb
ori
pT7
BstEII
lacI

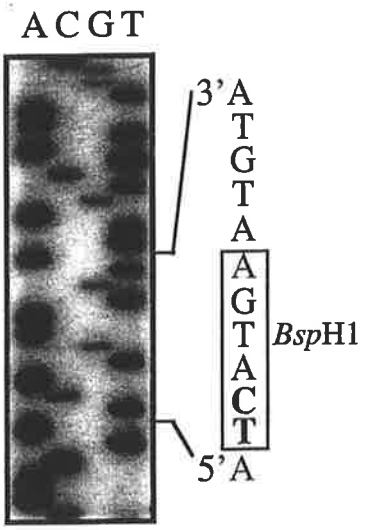
Figure 3.2: Nucleotide changes introduced to generate a bacterial expression vector for recombinant expression of yBPL-His.

Mutated plasmids were sequenced using the method described by (Sanger *et al.*, 1977) and the products resolved on 6% polyacrylamide gels. The DNA sequence is shown to the right of the gels with the nucleotide changes highlighted in bold. The amino acid sequence of the translated product is shown on the far right.

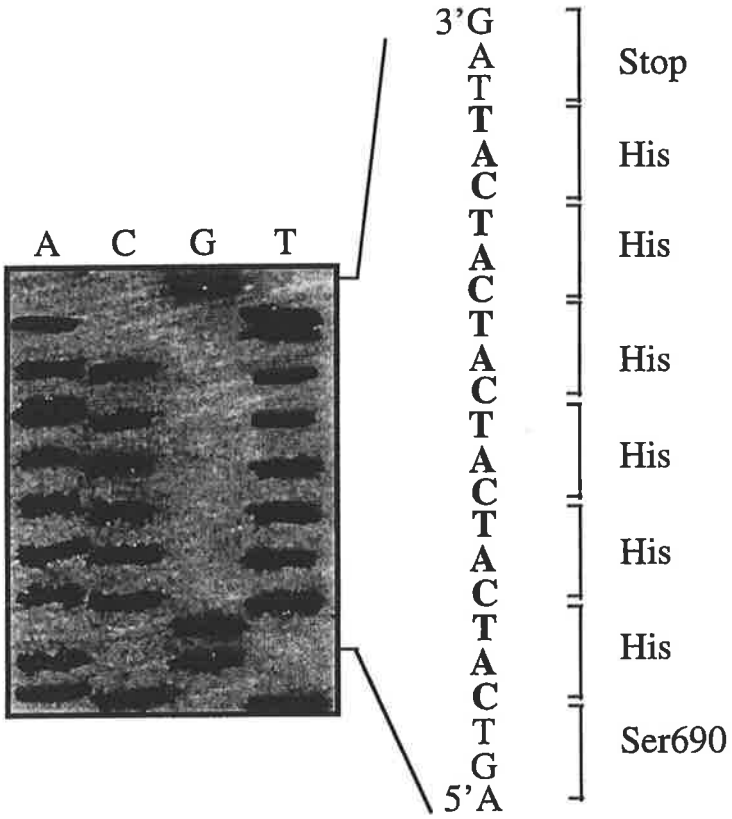
A) Introduction of a *Bsp*HI endonuclease restriction site (boxed) around the initiation of translation codon (Met1) facilitated cloning of the product into the pET expression vector.

B) The codons for six histidine residues were introduced into 3' region of the *bpl* coding region. The translated product fused the hexahistidine sequence onto the C-terminus (Ser690) of yBPL.

A


ACGT
3'A
T
G
T
A
A
G
T
A
C
T
5'A
BspH1
Val3
Asn2
Met1

B


A C G T
3'G
A
T
T
A
C
T
A
C
T
A
C
T
A
C
T
A
C
T
A
C
T
A
C
T
G
5'A
Stop
His
His
His
His
His
His
Ser690

expression of yBPL containing six histidine residues at the C-terminus. The C-terminus was chosen as the site for the His6-tag for two reasons. Firstly, the crystal structure of BirA revealed that the four C-terminal residues were poorly defined and assumes multiple conformations (Wilson *et al.*, 1992). Thus, if this region of yBPL is structurally analogous to BirA, the introduction of the His6-tag is unlikely to disrupt the structure or function of the enzyme. Secondly, as the proposed catalytic site of yBPL lies in the C-terminal region, N-terminally truncated homologues of the enzyme could be constructed and detected with C-terminal the His-tag. Therefore, six histidine codons were introduced at the 3' end of the gene by PCR mutagenesis using primers BplPst and BplBam with pKS(Bpl) as template (Figure 3.2B). The expression vector pET(Bpl-His) was constructed by a three fragment ligation containing the purified PCR product from this reaction which had been treated with *Pst*I / *Bam*HI, the 1.1 kb *Pst*I / *Bam*HI fragment from pET(Bpl) and *PstI*-digested pET16b.

3.3.2 Protein expression

The yBPL expression vectors constructed above were transformed into the engineered *E. coli* strain BL21(λDE3)pLysS (Novagen). Upon induction with IPTG this strain expresses T7 RNA polymerase which is under the control of the *lac* promoter. The bacteria also harboured the pLysS plasmid for expression of bifunctional T7 lysozyme. T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase, represses low level expression from the T7 promoter as well as permitting the lysis of cells by freeze thawing. Expression of recombinant yBPL and yBPL-His were driven by the T7 RNA polymerase as the *bpl* and *bpl-His* genes were under control of the T7 promoter. Induction of the recombinant

enzymes with IPTG was confirmed by analysis of crude lysates by SDS-PAGE (Figure 3.3A) An induced band of approximately 76 kDa was observed in those bacteria containing either of the two expression vectors but was not observed in the same cells harbouring the parent vector. Assays for yBPL activity, carried out on crude lysates from cells expressing yBPL with and without the His6-tag, revealed that the presence of the His6-tag did not affect the activity of the enzyme (Figure 3.3B). Therefore all subsequent expressions were performed using the tagged protein only. Initial experiments showed that growth at 37°C and induction with 1 mM IPTG led to inclusion body formation, in both shake flasks (Figure 3.4A) and a 1L fermenter. For the recovery of soluble enzyme from the protein expression system, cells were grown at 30°C in shake flasks and expression of yBPL was induced with 0.1 mM IPTG (Figure 3.4B).

3.3.3 Purification of yBPL

3.3.3.1 Nickel chelating chromatography

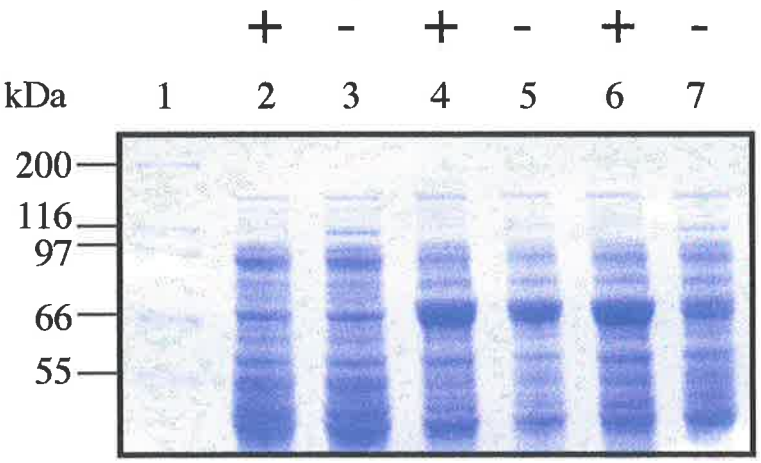
Proteins fused with a hexahistidine motif have been shown to interact specifically with immobilised metal ions at a pH greater than 6 where the histidine side chains are non-protonated. As an initial purification step, crude lysate containing yBPL-His was passed over a Ni-NTA column, giving a 30-40-fold enrichment (Figure 3.5A). Non-specific adhesion of metal binding proteins to the resin is reduced by the addition of imidazole to the loading and washing buffers. However yBPL-His6 eluted from the resin with only 20 mM imidazole. Therefore column washings were performed under low stringency conditions (10 mM imidazole) which allowed the non-specific binding of bacterial proteins. BPL

Figure 3.3 Expression and activity of recombinant yBPL and yBPL-His.

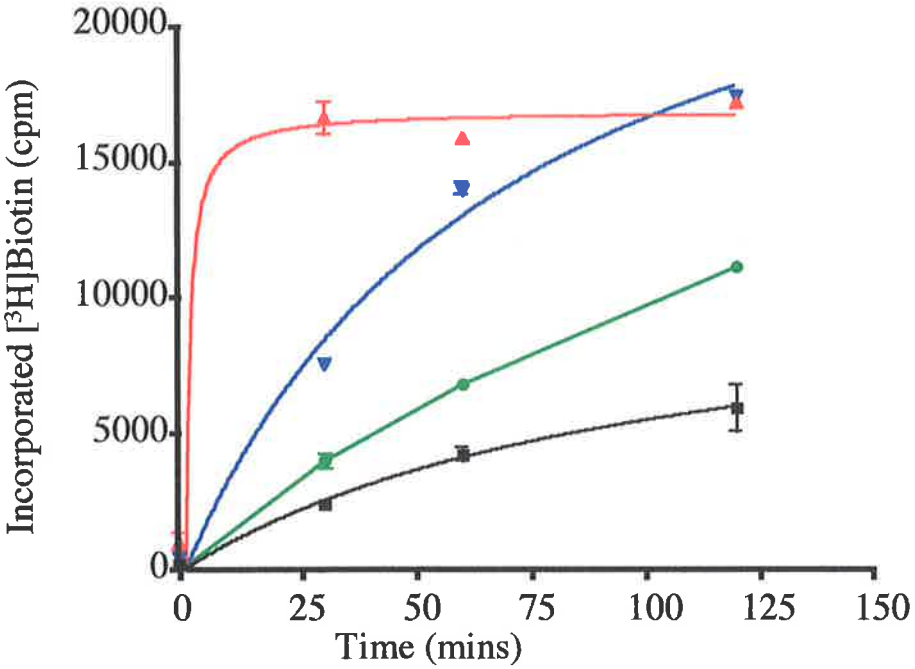
A) Expression of recombinant yBPL and yBPL-His was analysed by SDS-PAGE. Whole cell lysates were fractionated on 10% polyacrylamide gels under reducing conditions. Protein standards were loaded in lane 1 and their molecular masses are indicated to the left of the gel. Cells were grown at 37°C in the presence (+) or absence (-) of 1 mM IPTG. *E. coli* BL21(λDE3)pLys strains harbouring pET16b (lanes 2 & 3), pET(Bpl) (lanes 4 & 5) or pET(Bpl-His) (lanes 6 & 7) were analysed for protein induction.

B) The activity of BPLs in crude cell lysates was assayed by measuring the incorporation of [^{3}H]biotin onto TCA precipitable protein (3.2.3). The assay was performed over the time scale indicated on crude lysates (100 μg) containing overexpressed BirA (Chapman-Smith et al, 1994) (red line), BPL (green line), BPL-His (blue line) and endogenous BirA (black line).

A


+ - + - + -
kDa 1 2 3 4 5 6 7
200
116
97
66
55

B


Incorporated [3H]Biotin (cpm)
20000
15000
10000
5000
0
0 25 50 75 100 125 150
Time (mins)

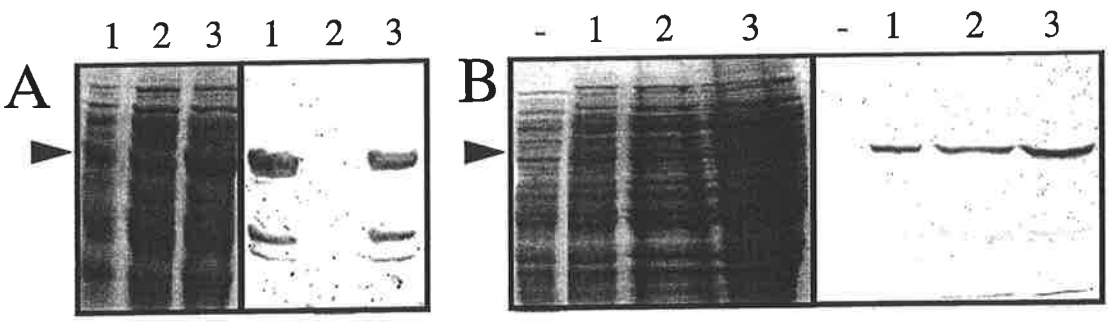


Figure 3.4: Solubility of yBPL-His under different induction conditions

Expression of recombinant yBPL-His in *E. coli* BL21(λDE3)pLys cells was performed under two different conditions. (A) Initially cells were induced with 1 mM IPTG at 37°C for 3 hours. (B) The induction was repeated with 0.1 mM IPTG at 30°C for 2 hours. Protein samples (50 µg) were fractionated on 10% polyacrylamide gels and analysed by coomassie blue staining (left panels) and NiNTA blots (right panels). Lanes were loaded with (1) whole cell lysate, (2) soluble material and (3) insoluble material. Lanes marked (-) represent uninduced cultures. Arrows indicate the migratory position of yBPL-His.

Figure 3.5 Purification of recombinant yBPL-His

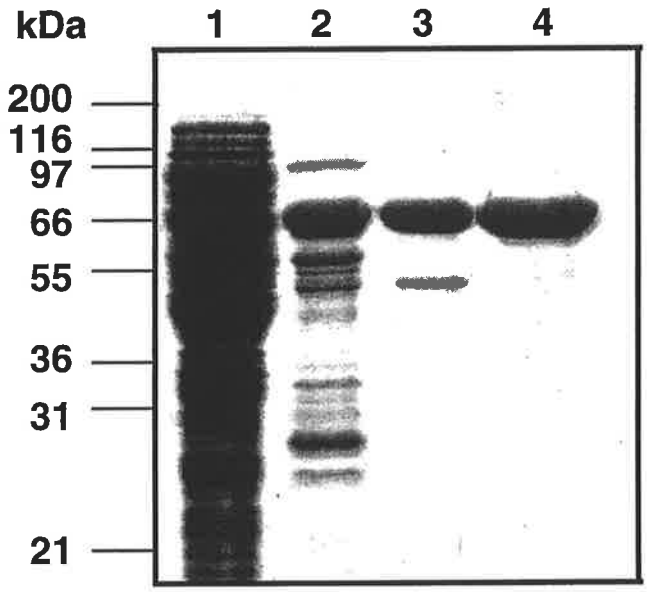
A) Purification of recombinant yBPL from a 2l culture was performed as described in 3.2.2. The data given is a representative example from three experiments. The Units shown here are nmol of holo protein formed per minute.

B) The purification of yBPL was analysed by SDS-PAGE. Fractions were separated on a 12% polyacrylamide gel under reducing conditions. Lane 1 represents 50 µg of crude extract, lane 2 represents 10 µg of material obtained from NiNTA chromatography, lane 3 represents 5 µg of material obtained from the Q-Sepharose step and lane 4 represents 5 µg of the protein obtained after gel filtration. The migration of molecular mass markers is indicated on the left.

A

Step	Protein (mg)	Total Units	Specific Activity (Units/mg)	Fold Purified	% Recovery
Crude	326	848	2.6	(1)	100
NiNTA	8.8	837	93.9	37	99
Q-Sepharose	1.75	400	228.6	88	47
Superdex-75	0.9	263	293	113	31

B



activity was detected in the unbound material, but as no 77 kDa protein containing a His6-tag was detected in this fraction by Ni-NTA blot analysis, the activity apparently represented endogenous bacterial biotin ligase. The recombinant yBPL obtained by Ni-NTA chromatography was immediately dialysed overnight against storage buffer, since storage in imidazole resulted in irreversible inactivation of the enzyme.

3.3.3.2 Ion Exchange

Additional purification using ion exchange chromatography was required following Ni-NTA chromatography, as the low stringency washing conditions failed to remove a number of contaminating proteins (Figure 3.5B, lane 2). A predicted pI of 6.6 for yBPL-His, based on the primary structure of the protein, was determined using a program available through the Swiss-Prot database. Anion exchange was selected as adhesion of the enzyme to Q-Sepharose resin could be performed at a close to neutral pH. The enzyme eluted from the column with 125 mM NaCl (Figure 3.6) and fractions containing material which reacted with the NiNTA conjugate by blot detection were pooled. After two chromatography steps, half of the initial enzyme activity had been recovered, and the protein was at least 80% pure. This represented a purification of around 88-fold (Figure 3.5A). However, in addition to the 77 kDa protein corresponding to full length His6-yBPL, a band of 50 kDa was observed on SDS-PAGE of the pooled fractions from the Q-Sepharose column (Figure 3.5B, lane 3). This protein was also detected on a Ni-NTA blot, indicating the presence of the C-terminal His6-tag. N-terminal sequencing showed the 50 kDa protein to be a proteolysis

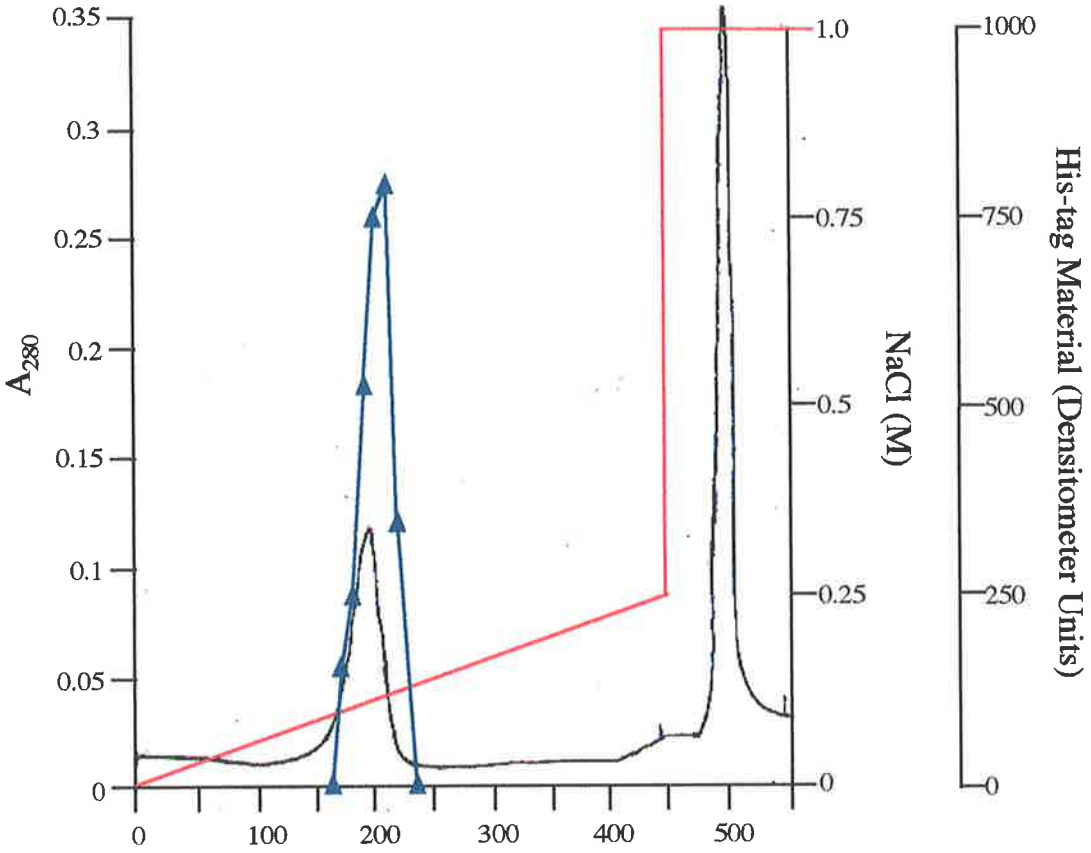


Figure 3.6 Ion exchange purification of yBPL-His

Protein obtained from NiNTA chromatography was further purified using Q-Sepharose chromatography as described in 3.2.2. Protein eluted from the column was monitored by absorbance at 280 nm (—). A gradient of 0 to 250 mM NaCl over 450 millilitres (—) was employed to elute protein off the column. Fractions containing His-6 tagged material were determined by NiNTA blot and the intensity of the bands quantitated using laser densitometery (▲—▲).

product of the intact yBPL, with cleavage occurring between Lys-248 and Phe-249.

3.3.3.3 Gel Filtration

In order to obtain a purified sample of full length yBPL-His, a method was required which resolved the cleaved and intact forms of the enzyme obtained after ion exchange. Since these proteins also coeluted during hydrophobic interaction chromatography on Phenyl Sepharose6 resin, another means of purifying yBPL-His was required. Investigation of this material by analytical gel filtration on both Superdex 75 or 200 columns revealed that the 77 kDa and 50 kDa proteins could only be partially resolved. In order to determine if the cleaved 50 kDa fragment still associated with the N-terminal fragment, and therefore had the characteristics of a larger molecule, material obtained after Ni-NTA and ion exchange purification was analysed using a Superdex 75 column. This column was able to resolve 66 kDa bovine serum albumin from 30 kDa carbonic anhydrase (Figure 3.7A). Analysis by SDS-PAGE of the gel filtration fractionated yBPL-His sample demonstrated that the cleavage products did not remain associated after 3 purification steps, since a product around 27 kDa fragment was not found in stoichiometric amounts in fractions containing the 50 kDa fragment (Figure 3.7B).

Protein collected after anion exchange chromatography was concentrated by ultrafiltration prior to fractionation on a preparative Superdex 75 gel filtration column (26 x 600 mm; Pharmacia) run in Storage Buffer. As the quantity of cleaved yBPL constituted less than 20% of the sample (as determined by laser densitometry of bands detected on SDS-PAGE), approximately 50% of the intact

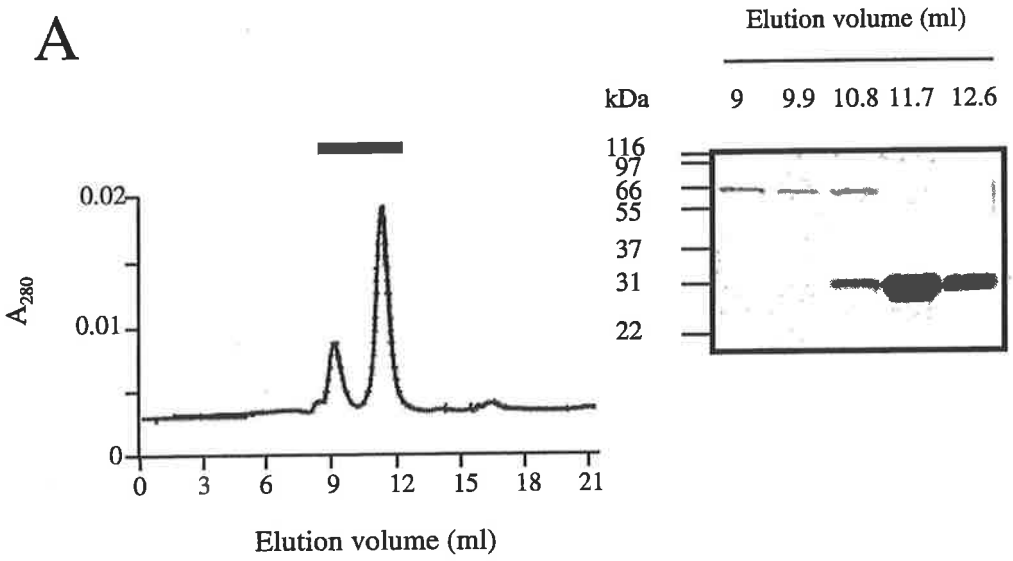
Figure 3.7 Analytical Gel Filtration

Recombinant yBPL-His and molecular weight standards were analysed by gel filtration chromatography using a Superdex-75 HR 10/30 column equilibrated and run under native conditions in Storage Buffer. Material eluted from the column was detected by absorbance at 280 nm and analysed by reducing SDS-PAGE (gel shown to right of chromatogram) with molecular weight markers together with molecular mass markers. The bar above the chromatograms indicates the eluted material collected for analysis by SDS-PAGE.

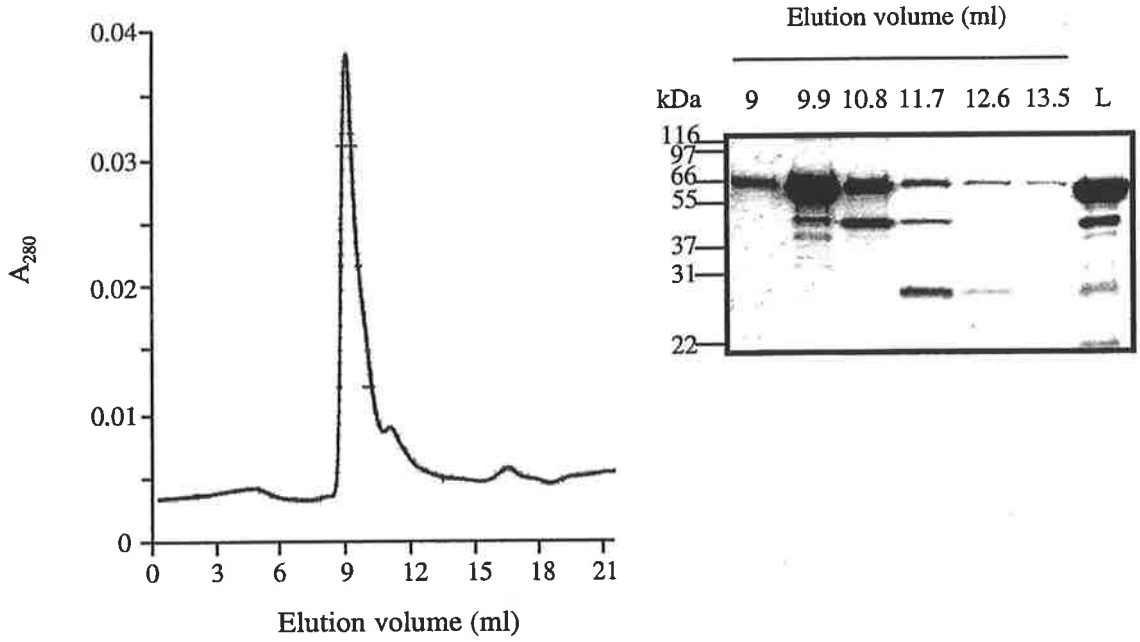
A) 100 µg of bovine serum albumin (Mr 66 000) was separated from 100 µg of bovine carbonic anhydrase (Mr 29 000).

B) A sample (230 µg) of ion exchange purified BPL-His was fractionated by gel filtration collecting 0.9 ml fractions. 100 µl of fractions 10 - 15 were analysed by SDS-PAGE. The lane marked L represents 10 µg of the starting sample.

A


Elution volume (ml)
kDa
9
9.9
10.8
11.7
12.6
116
97
66
55
37
31
22
0.02
0.01
0
A_{280}
0
3
6
9
12
15
18
21
Elution volume (ml)

B


Elution volume (ml)
kDa
9
9.9
10.8
11.7
12.6
13.5
L
116
97
66
55
37
31
22
0.04
0.03
0.02
0.01
0
A_{280}
0
3
6
9
12
15
18
21
Elution volume (ml)

enzyme in the sample was resolved away from the cleaved product by collecting only the leading fractions of the protein peak. These fractions contained full length yBPL with a purity of greater than 95%, determined by SDS-PAGE (Figure 3.5, lane 4) and N-terminal sequencing. The intact yBPL purified in this way was used for all subsequent kinetic and proteolysis analysis.

3.3.4 Conclusions

While BPL has been purified from a variety of sources, the low abundance of the endogenous enzyme has made its purification a difficult task. Since the availability of recombinant DNA technology has permitted high level production of proteins in suitable hosts with improved yields, protein overexpression has facilitated isolation of BirA (Eisenberg *et al.*, 1982; Buoncristiani & Otsuka, 1988) and *Arabidopsis thaliana* BPL from *E. coli*. Here recombinant expression in bacteria of a large, eukaryotic member of this enzyme family was performed. Cloning a hexahistidine tag onto the C-terminus of yBPL permitted isolation of active enzyme in a rapid two step purification, with yields comparable to those reported for recombinant production of the bacterial protein (Buoncristiani & Otsuka, 1988; Eisenberg *et al.*, 1982). Whereas partial purification of BPL from *S. cerevisiae* has been reported previously (Sundaram *et al.*, 1971), our system allowed purification to apparent homogeneity (Figure 3.5).

The use of gel filtration as a means of resolving the cleaved and intact yBPL is an inefficient step for large scale purification of the enzyme. An alternative, final chromatography step in the protocol to replace gel filtration could be investigated. However, a number of techniques have been attempted and finding an optimal step would be time consuming. One method of improving

the yields of recombinant yBPL is to produce an analogue which is resistant to proteolysis in the bacteria. Endogenous cleavage of the protein was shown to occur between Lys 248 and Phe 249, where the basic side chain of lysine is most likely required for protease recognition. Thus, substitution of Lys 248 with a conservative but non-charged residue should reduce *in vivo* proteolysis and, substitution of Lys 248 for Thr would be the most appropriate change (Bordo & Argos, 1991). Providing the mutated enzyme retains yBPL activity, this variant may be more suitable for large scale purification.

Chapter 4

KINETIC ANALYSIS OF RECOMBINANT YEAST BIOTIN PROTEIN LIGASE

KINETIC ANALYSIS OF RECOMBINANT YEAST BIOTIN PROTEIN LIGASE

4.1 INTRODUCTION

In the previous chapter, the purification of recombinant yeast biotin ligase to near homogeneity was described. Before using this material for detailed kinetic analysis, the assay conditions required for optimal activity needed to be determined. BPLs from a wide variety of sources have been purified and investigated. Although the members of this enzyme family perform the same essential reaction, the assay conditions are not the same for all members. Factors such as optimal pH, the inclusion of salts, the specificity of metal ions and the nucleotide triphosphate source can differ for each enzyme and in each case require investigation.

The reaction mechanism catalysed by BPLs appears to be a Bi Uni Uni Bi ping pong mechanism, that is a double displacement kinetic mechanism (Tissot *et al.*, 1998). This model implies the formation of an enzyme-bound biotinyl-5'-AMP intermediate in the first partial reaction from biotin and ATP. The order which the substrates bind to the enzyme in the kinetic reaction mechanism is intriguing. For BirA biotin is the first substrate to bind, whereas ATP binding precedes the introduction of biotin into the kinetic mechanism catalysed by plant BPL (Tissot *et al.*, 1998). As yet, a detailed kinetic analysis of this manner has not been performed using a large eukaryotic BPL, such as from yeast or humans. An understanding of the reaction mechanism is important for evaluating existing or for improving therapies against defective BPLs in MCD.

4.2 RESULTS

4.2.1 Biological properties of yBPL

The activity of yBPL was investigated using steady state kinetics by assaying the velocity of the enzyme within the first 10% of the reaction. yBPL activity was determined by measuring the incorporation of [^{3}H]biotin into one of two biotin accepting domains; either the C-terminal 87 amino acid residues of the *E. coli* biotin carboxyl carrier protein (BCCP-87) (Chapman-Smith *et al.*, 1999) or the C-terminal 104 amino acid residues of *S. cerevisiae* pyruvate carboxylase 1 (yPC-104) (Val *et al.*, 1995). The presence of magnesium ions, ATP, biotin and the apo form of a biotin domain were necessary substrates for activity (Table 4.1). The inclusion of EDTA further inhibited the reaction, suggesting the presence of a divalent metal ion bound to the enzyme (Table 4.1). Initially, the optimal conditions for the BirA reaction were tested for yBPL. Systematic investigation of each reaction component was then performed to determine the conditions required for optimal activity for yBPL.

Table 4.1: Removal of reaction components on yBPL activity

Reaction Conditions	Activity (pmol/min)	Relative Activity
Complete	0.962	100%
- ATP	0.017	1.8%
- $MgCl_2$	0.069	7%
- $MgCl_2$ + 1mM EDTA	0.013	1.4%
- Biotin	0	0%
- Apo BCCP-87	0	0%
- Ligase	0	0%

4.2.1.1 Optimal pH

Optimal yBPL activity was investigated using two buffer systems; Tris-HCl (pH 7.1 - 8.9) and NaH_2PO_4-Na_2HPO_4 (pH 5.8 - 8.0). In the physiological pH range between pH 7.0 - 8.0, yBPL displayed the greatest activity with Tris-HCl. Using the Tris-HCl buffer, optimal enzyme activity was observed between pH 8.0 and 8.5.

4.2.1.2 Divalent and Monovalent Metal Ions

A range of divalent metal ions was investigated to see which could be substituted for magnesium in the reaction. As shown in Table 4.2, calcium, nickel and manganese ions supported activity to varying extents. However cobalt, zinc, and copper ions were a poor substitute for magnesium. Interestingly, monovalent rubidium ions were able to sustain complete enzyme activity. For the BirA reaction, 100 - 200 mM potassium ions are included (Beckett & Matthews, 1997; Chapman-Smith *et al.*, 1999). In contrast, the presence of 200 mM KCl inhibited yBPL activity at pH 7.5 and pH 8.0. (Figure 4.1).

Table 4.2: Activity of yBPL with 5.5 mM of various metal ions

Metal Ion	Activity (pmol/min)	Relative Activity
$MgCl_2$	0.442 ± 0.19	100%
$CaCl_2$	0.483 ± .0138	109.4%
RbCl	0.468 ± 0.182	106%
$NiCl_2$	0.527 ± .033	74%
$MnCl_2$	0.236 ± 0.11	53.3%
$CoCl_2$	0.063 ± .05	14.2%
$ZnCl_2$	0.006 ± .0004	1.3%
$CuCl_2$	0.002 ± 0.005	0.45%

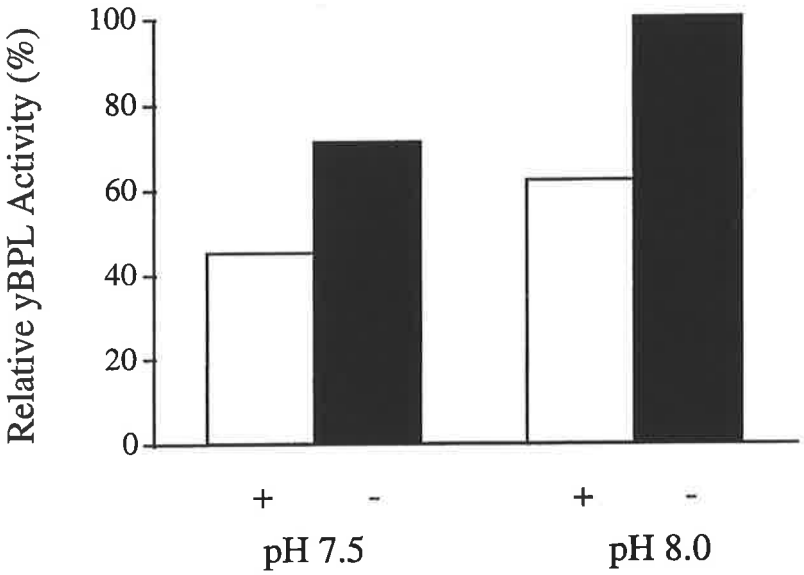


Figure 4.1: Activity of yBPL in the presence of salt.

The activity of yBPL was assayed at pH 7.5 and pH 8.0 in the absence (-) and presence (+) of 200 mM KCl. The activity of the enzyme under these conditions was expressed relative to the activity observed at pH 8.0 without KCl.

4.2.1.4 Nucleotide Triphosphate

yBPL activity was very highly dependent on ATP as a nucleotide triphosphate source. Substitution of ATP with CTP, GTP and UTP gave 6% or less of the activity observed when ATP was the NTP source (Table 4.3).

Table 4.3: Activity of yBPL with various nucleotide triphosphates

Nucleotide Triphosphate	Activity (pmol/min)	Relative Activity
ATP	0.782 ± 0.177	100%
UTP	0.046 ± 0.001	5.9%
GTP	0.005 ± 0.003	0.6%
CTP	0.003 ± 0.001	0.4%

4.2.1.5 Specificity for biotin

The ability of the biotin analogues biocytin, diaminobiotin, desthiobiotin and iminobiotin to inhibit the yBPL reaction was investigated. At a concentration of 5 μM, only unlabelled biotin was able to inhibit the incorporation of 50 nM [^{3}H]biotin into acceptor protein. Furthermore, addition of lipoic acid (5 μM) or acetyl CoA (5 μM) had no inhibitory effect on activity.

4.2.2 Kinetic analysis of yBPL

4.2.2.1 Determination of K_m Values

The kinetic constants for d-biotin, MgATP and two different biotin domain substrates, apo BCCP-87 and apo yPC-104 were determined using steady state

kinetics (Figure 4.2). The K_m for MgATP was determined to be 20.9 ± 3 μM (Figure 4.2A). As has been observed with BPLs from a wide variety of species, the yeast enzyme has a low K_m for biotin; 67 ± 11 nM (Figure 4.2B). The yeast substrate, apo yPC-104, displayed a low K_m of 1.0 ± 0.2 μM (Figure 4.2C) whereas the K_m for the bacterial biotin domain was more than 10-fold higher (11.1 ± 1μM; Figure 4.2D).

4.2.2.2 Order of Substrate Binding

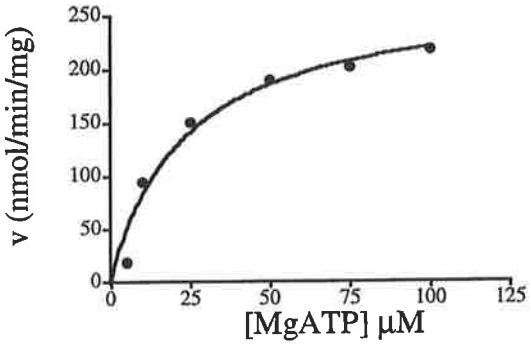
In order to determine the order of substrate addition in the kinetic pathway, the activity of yBPL was assayed under steady state conditions when the concentration of two substrates were varied while maintaining the third at saturating levels. The double-reciprocal plots from the velocity measurements are shown in Figure 4.3. When biotin and MgATP were the varied substrates patterns of intersecting lines were obtained (Figure 4.3A & B), indicating a reversible connection between these two substrates in the reaction pathway (Cleland, 1970). When the concentration of acceptor protein was varied together with either biotin or MgATP, patterns of parallel lines were observed on the double-reciprocal plots (Figure 4.3C & D). These results implied both biotin and MgATP combine with yBPL before the acceptor protein (Cleland, 1970). Biotinyl-5'-AMP is known to react readily with hydroxylamine to form biotinyl-hydroxamate (Christner *et al.*, 1964; Morita *et al.*, 1998). When yBPL was incubated with [^{3}H]biotin, MgATP and hydroxylamine, [^{3}H]biotinyl-hydroxamate was formed confirming that the reaction pathway proceeds through the formation of the biotinyl-5'-AMP intermediate.

Figure 4.2: Steady-state kinetic analysis of yBPL substrate binding

The activity of yBPL-His was measured under steady state conditions which kept two substrates at constant, saturating levels while varying the concentration of the third substrate over the ranges indicated on the graphs. From the curves, K_m values for A) MgATP, B) biotin, C) apo yPC-104 and D) apo BCCP-87 were determined. The lines represent non-linear regression to the Michaelis-Menten equation using GraphPad Prism, as described in 3.2.3.

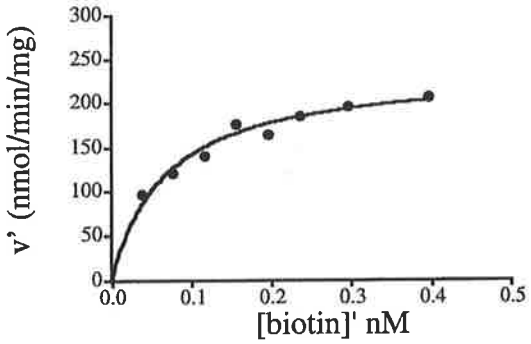
The saturating concentration of substrates employed in the assays were 3 mM ATP, 5 μM biotin and 5 μM apo yPC-104.

A



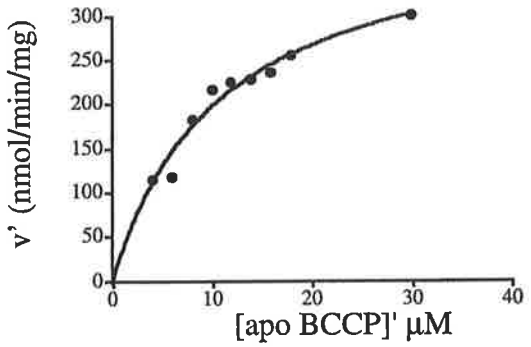
$K_m = 20.9 \pm 3$ μM

B



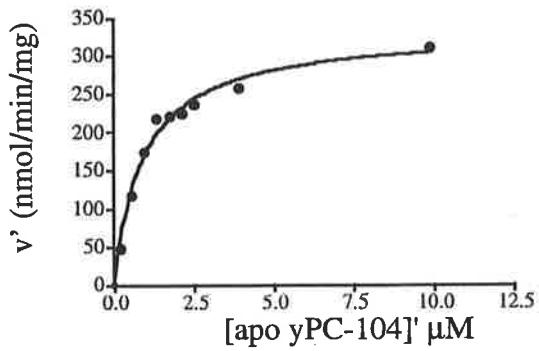
$K_m = 67 \pm 11$ nM

C



$K_m = 11.1 \pm 1$ μM

D



$K_m = 1.0 \pm 0.2$ μM

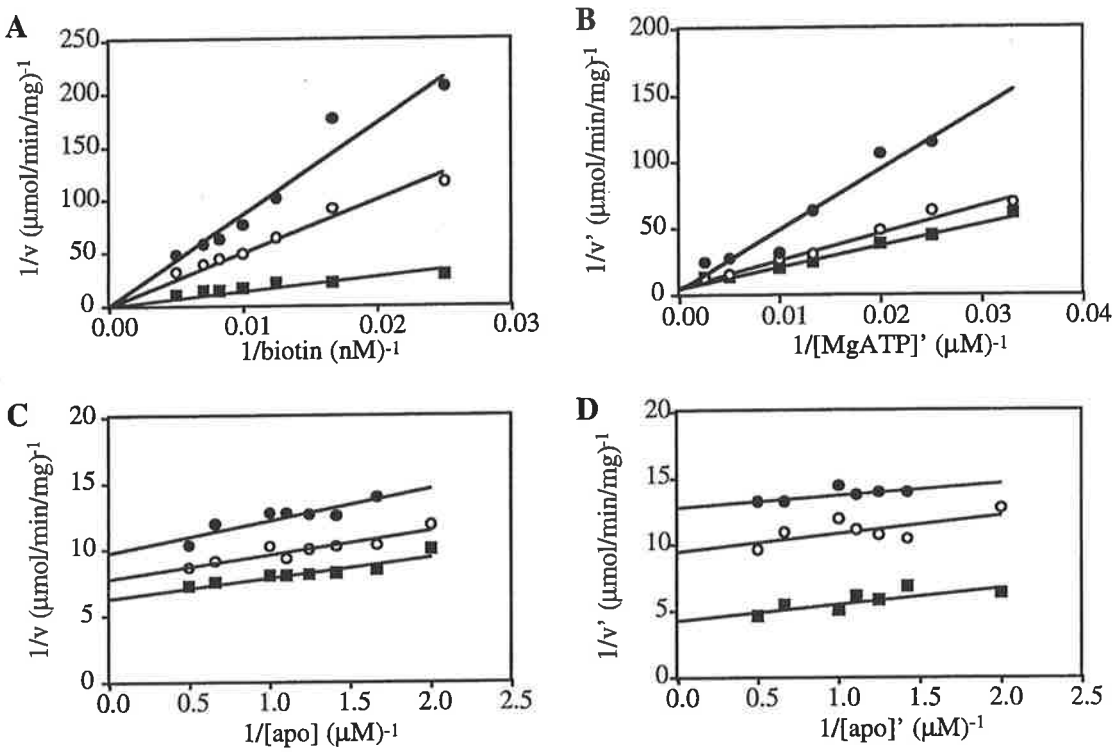


Figure 4.3: Steady-state kinetic analysis of yBPL ordered substrate binding

Double reciprocal plots of initial velocities with A) variable biotin concentrations and different fixed concentrations of MgATP, 50 µM (●), 100 µM (O) and 500 µM (■); B) variable MgATP concentrations and different fixed concentrations of biotin, 25 nM (●), 50 nM (O) and 70 nM (■); C) variable apo yPC-104 concentrations and different fixed concentrations of biotin, 25 nM (●), 35 nM (O) and 50 nM (■) and D) variable apo yPC-104 concentrations and the different fixed concentrations of MgATP as in A). Other assay conditions were as described in 3.2.3.

4.2.2.3 Inhibition by Pyrophosphate

As the formation of biotinyl-5'-AMP is accompanied by the release of pyrophosphate in the reaction mechanism (Eq 2), product inhibition studies with pyrophosphate were performed. Pyrophosphate behaved as a competitive inhibitor relative to MgATP and as a non-competitive inhibitor with respect to biotin (Figure 4.4) implying that MgATP binding precedes biotin binding (Cleland, 1970).

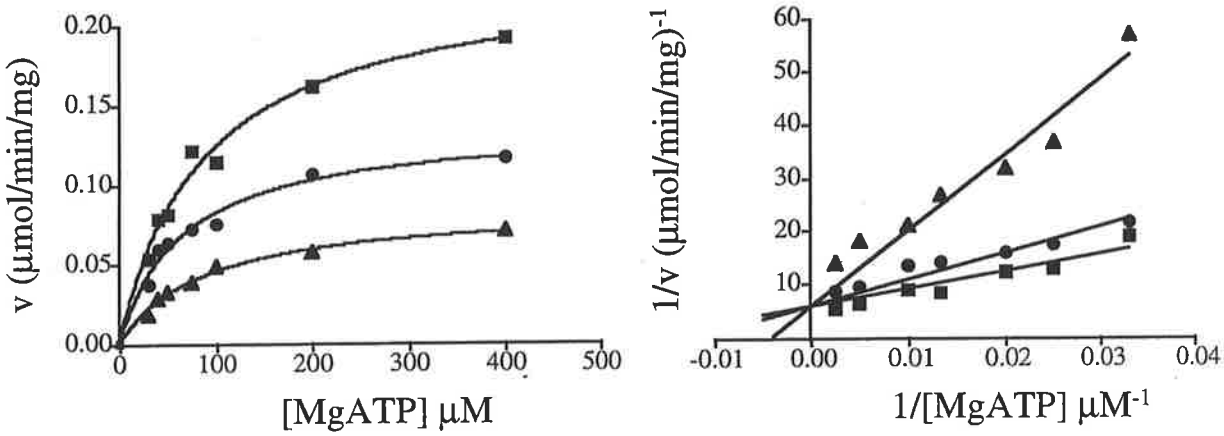
Figure 4.4: Inhibition of yBPL activity with pyrophosphate

Inhibition of yBPL activity by pyrophosphate was subjected to Lineweaver-Burk analysis. Concentrations of pyrophosphate included in the reactions were 0 μM (■), 50 μM (▼), 100 μM (●) and 500 μM (▲).

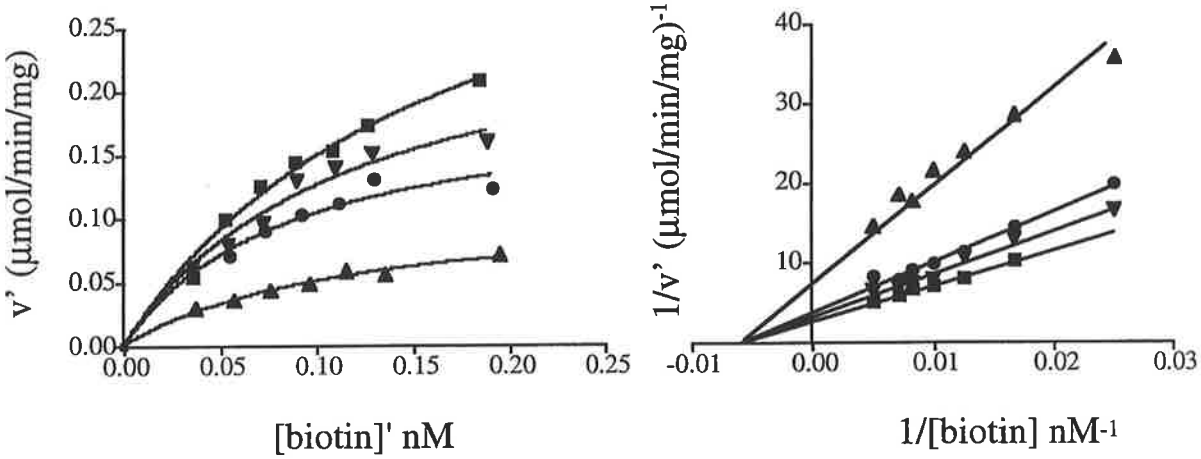
A) Double reciprocal plots of initial velocity with varying MgATP and different fixed concentrations of pyrophosphate.

B) Double reciprocal plots of initial velocity with varying biotin concentrations and different fixed concentrations of pyrophosphate. The ATP concentration was 0.5 mM. Other assay conditions were as described in 3.2.3.

A


v (µmol/min/mg)
0.20
0.15
0.10
0.05
0.00
0
100
200
300
400
500
[MgATP] µM
1/v (µmol/min/mg)-1
60
50
40
30
20
10
-0.01
0.00
0.01
0.02
0.03
0.04
1/[MgATP] µM-1

B


v' (µmol/min/mg)
0.25
0.20
0.15
0.10
0.05
0.00
0.00
0.05
0.10
0.15
0.20
0.25
[biotin]' nM
1/v' (µmol/min/mg)-1
40
30
20
10
-0.01
0.00
0.01
0.02
0.03
1/[biotin] nM-1

4.3 DISCUSSION

The results of steady-state kinetic analysis performed upon yBPL indicate that biotinylation occurs through a two step Bi Uni Uni Bi ping pong mechanism (Eq 2 and 3). In the first partial reaction, the enzyme complexes with ATP and biotin and catalyses the synthesis of biotinyl-5'-AMP with subsequent release of pyrophosphate. The addition of the apo acceptor protein then follows in the second partial reaction with the release of the biotinylated protein and AMP. The kinetics of the formation of the adenylated intermediate by BirA has been analysed quantitatively (Xu & Beckett, 1994). The BirA-biotinyl-5'-AMP complex is quite stable (Xu & Beckett, 1994) and is proposed to be the most abundant enzyme form in the cell. Here we demonstrate that the precursors of the intermediate, ATP and biotin, bind to yBPL in an ordered manner. As has been observed with plant BPL (Tissot *et al.*, 1998), ATP binds to the yeast enzyme before biotin. This is in contrast to the reaction pathway catalysed by bacterial BirA where biotin is the first ligand to bind the enzyme (Xu & Beckett, 1994). The order of substrate binding in *E. coli* is believed to allow more responsive regulation of biotin biosynthesis. When the cellular demand for biotin is low, the BirA-biotinyl-5'-AMP complex occupies the *bio* operator sequence and represses transcription of the biotin biosynthetic operon. As apo BCCP levels increase, biotin is transferred from the enzyme-bound adenylate to the protein bound form, with concomitant derepression of the biotin operon. Kinetic analysis of the interaction of BirA with biotin and ATP indicates that formation of the repressor complex is highly sensitive to biotin, and the K_m for ATP is in the low millimolar range (Xu & Beckett, 1994; Beckett & Matthews, 1997; Chapman-Smith *et al.*, 1999). In contrast, BPLs from biotin auxotrophic species generally bind ATP at

lower concentrations, with K_m values in the range 0.38 μM to 200 μM (Siegel *et al.*, 1965; Murthy & Mistry, 1974; Chiba *et al.*, 1994; Xia *et al.*, 1994). The value of 21 μM reported here for *S. cerevisiae* is well below the intracellular concentration of ATP (Wallace *et al.*, 1994) and is consistent with the absence of repressor function in the eukaryotic BPLs.

The formation of an activated enzyme-bound biotinyl intermediate in the first partial reaction of BPL requires a nucleotide triphosphate and a divalent metal ion (Christner *et al.*, 1964). BPLs from different organisms differ in their specificity for both the NTP source and the divalent metal. For example, magnesium is the preferred metal ion for pea BPL (Tissot *et al.*, 1996) whereas zinc and manganese ions can readily substitute for magnesium ions for the *P. shermanii* enzyme (Lane *et al.*, 1964; Shenoy & Wood, 1988). For the yeast enzyme, calcium or magnesium ions were the preferred divalent metals, with reasonable levels of activity also seen in the presence of nickel or manganese, while cobalt, zinc, and copper ions failed to support significant activity. Yeast BPL essentially had an absolute requirement for ATP, as only minimal enzyme activity was detected when ATP was replaced by other nucleotide triphosphates. This is similar to BPLs from *P. shermanii*, *B. stearothermophilus*, pig and rabbit which are also specific for ATP (Christner *et al.*, 1964; Siegel *et al.*, 1965; Cazzulo *et al.*, 1971; Shenoy & Wood, 1988). However, UTP can replace ATP for the enzyme from chicken liver (Madappally & Mistry, 1970; Murthy & Mistry, 1974), whereas CTP is the preferred nucleotide for both the pea (Tissot *et al.*, 1996) and bovine liver (Chiba *et al.*, 1994) biotin ligases.

Kinetic studies on BPLs have reported a wide range of K_m values for biotin, ranging from 4.7 nM for the rabbit liver enzyme (Siegel *et al.*, 1965) to 3.3

μM in chicken liver (Murthy & Mistry, 1974). Here a low K_m for biotin (67 nM) for yeast BPL was determined. Baker's yeast is auxotrophic for biotin, which is actively transported into the cell via the recently cloned H^+-biotin symporter, VHT1 (Stolz *et al.*, 1999). This membrane bound transport protein displays maximal activity when cells are grown in media containing an extremely low concentration of biotin, 0.8 nM, and is inhibited by greater than 8 nM biotin (Stolz *et al.*, 1999). Rogers and Lichstein (1969a) demonstrated that the cellular concentration of free biotin in yeast can be increased against a concentration gradient. Under conditions which give maximal growth, the free intracellular biotin pool can reach a concentration of 70 μM (Rogers & Lichstein, 1969a) which would be saturating for yBPL. The mechanism of transport inhibition is not understood and the possibility that yBPL interacts with the receptor to regulate biotin metabolism in yeast is both speculative and interesting.

An upper limit for the bimolecular rate constant for the formation of the enzyme-biotin complex can be determined using the values obtained in the kinetic analysis. The calculated k_{cat}/K_m value for biotin of $6.0 \pm 0.08 \times 10^6$ $M^{-1}s^{-1}$ is several orders of magnitude smaller than that predicted for a diffusion controlled process. This value is 3-fold smaller than that of BirA (Xu & Beckett, 1997) but 10-fold larger than the value determined for *A. thaliana* BPL (Tissot *et al.*, 1998), suggesting subtle structural differences in the active sites of these BPLs. Biotin binding was determined to be highly specific for yBPL as several closely related biotin analogues and lipoic acid, as well acetyl CoA, all failed to inhibit the incorporation of biotin even when present at 100-fold excess over [^{3}H]biotin. This substrate specificity is a common feature of BPLs from a variety of sources (Lane *et al.*, 1964; Chiba *et al.*, 1994; Tissot *et al.*, 1996). Thus, the decreased activity of

biotin-dependent enzymes in rat liver seen after administration of lipoic acid (Zempleni *et al.*, 1997) seems unlikely to be due to a direct effect of lipoic acid on BPL *in vivo* . Similarly, Cazzulo and coworkers (1971) reported the activity of yBPL was improved in the presence of acetyl CoA. In their assay system the synthesis of PC activity from apo enzyme preparations was measured. It has been previously reported that yeast PC is allosterically activated by the coenzyme (Cazzulo & Stopani, 1968). Together the data suggest that BPL is not likely to be activated by acetyl CoA.

We have used two isolated domains as the biotin acceptor domain in our analysis of yBPL, yPC-104 and BCCP-87. This latter peptide has been shown to be as effective a substrate for biotinylation by BirA as intact BCCP (Nenortas & Beckett, 1996), with a K_m of 4 μM in our assay system (Chapman-Smith *et al.*, 1999). In the present study a similarly low K_m value was determined when a yeast biotin domain was assayed with yBPL (1 μM). However, we observed a greater than 10-fold higher K_m (11 μM) when BCCP-87 was the substrate for yBPL. There is evidence of cross-species reactivity in biotinylation reactions (McAllister & Coon, 1966; Cronan, 1990; Leon-Del-Rio *et al.*, 1995) but kinetic analyses of the interactions have not been previously performed. It seems likely that the differences in K_m for the two acceptor proteins for yBPL reflects subtle changes in substrate recognition or efficiency of biotin transfer between the two proteins in the assay system.

Chapter 5

DOMAIN MAPPING AND TRUNCATION ANALYSIS OF YEAST BIOTIN PROTEIN LIGASE

DOMAIN MAPPING AND TRUNCATION ANALYSIS OF YEAST BIOTIN PROTEIN LIGASE

5.1 INTRODUCTION

In Chapter 3, the purification of yBPL was described. Interestingly, the enzyme was relatively resistant to proteolysis in the bacteria with most of the recombinant protein being the full length protein. Only one major proteolytic product was detected and the cleavage event was determined to occur between residues 247 and 248. Protease sensitive regions of BirA have been characterised (Xu *et al.*, 1995) and are present in poorly defined, solvent exposed loops in the structure (Wilson *et al.*, 1992). These loops are in or near the catalytic domain of the enzyme and several residues in the loops become more defined with the addition of biotin or biotinyl-5'-AMP. In addition, determination of the structure of BirA has also revealed the presence of three domains; an N-terminal DNA-binding domain, the central catalytic domain and a C-terminal domain of unknown function. To date, only the structure of BirA has been determined. Furthermore, there is a lack of understanding of the domain structures of any of the eukaryotic BPLs and the regions of these molecules which are necessary for activity.

5.2 SPECIFIC METHODS

5.2.1 Construction of arabinose-inducible yBPL expression vectors

A vector was constructed to allow arabinose-inducible expression of full length yBPL in *E. coli*. This vector was subsequently employed for complementation assays conditionally lethal strains of *E. coli*. The expression vector pAra13 (Cagnon *et al.*, 1991) was chosen as recombinant proteins can be induced in any suitable bacterial strain from the arabinose promoter. The gene for the protein which regulates this promoter, araC, is also present in pAra13. Recombinant yBPL was expressed with the C-terminal hexahistidine (His6) tag, in order to detect expression of the enzyme. The cloning of yBPL is described below and is outlined in Figure 5.1A. The 2.1 kb *Nco*I / *Bam*HI fragment from pET(Bpl-His) (Chapter 3.3.1) was cloned into *Nco*I / *Bgl*II treated pAra13 (Cagnon *et al.*, 1991) yielding pN/B-Bpl-His. This construct was digested with *Nco*I and *Bst*EII and ligated to the 308 bp *Bsp*HI / *Bst*EII fragment liberated from pGEM(5'bpl) (Chapter 3.3.1), to yield p[Met^1]Bpl 1-690-His6.

Vectors for expression of N-terminally truncated forms of yBPL (Figure 5.1B) were constructed by introducing translation initiation codons, which also contained an *Nco*I cloning site, preceding residues 233, 369 and 409 of the full-length protein (Figure 5.2). Oligonucleotides Bpl233, Bpl369 or Bpl409 were included in separate polymerase chain reactions with oligonucleotide BplPstBack and p(Bpl)KS- as the template. The purified PCR products were digested with *Nco*I and *Pst*I and cloned into similarly treated pN/B-Bpl-His. This gave constructs pAra[Met^1]Bpl 233-690-His6, pAra[Met^1]Bpl 369-690-His6 and pAra[Met^1]Bpl 409-690-His6 respectively. All constructs were confirmed by DNA sequencing using oligonucleotides AraFor and AraFor2.

Figure 5.1: Construction of vectors for the expression of N-terminally truncated variants of recombinant yBPL-His

The strategy for constructing a series of vectors for the arabinose-induced expression of N-terminally truncated forms of yBPL is described in 5.2.1 and outlined here.

A) The plasmid pAra(Bpl-His) was constructed using a two step cloning strategy by introducing DNA encoding full length BPL-His into the parent vector pAra13.

B) The plasmid pN/B-Bpl-His, obtained during the cloning of pAra(Bpl-His), was employed as the parent vector for the construction of a series of vectors for expression of N-terminally truncated forms of BPL-His. The DNA present between the *Nco*I and *Pst*I restriction sites was excised and replaced by a series of PCR products treated with the same endonucleases. The sizes of the PCR products, cloned into the treated vector, are shown to the left.

Abbreviations: ori, bacterial origin of replication; Amp R, ampicillin resistance gene; lacI, lactose operon represser protein; paraB, promoter AraB gene; araC, arabinose regulatory protein. The position of restriction sites employed in the cloning strategy are shown on the outside of the plasmids.

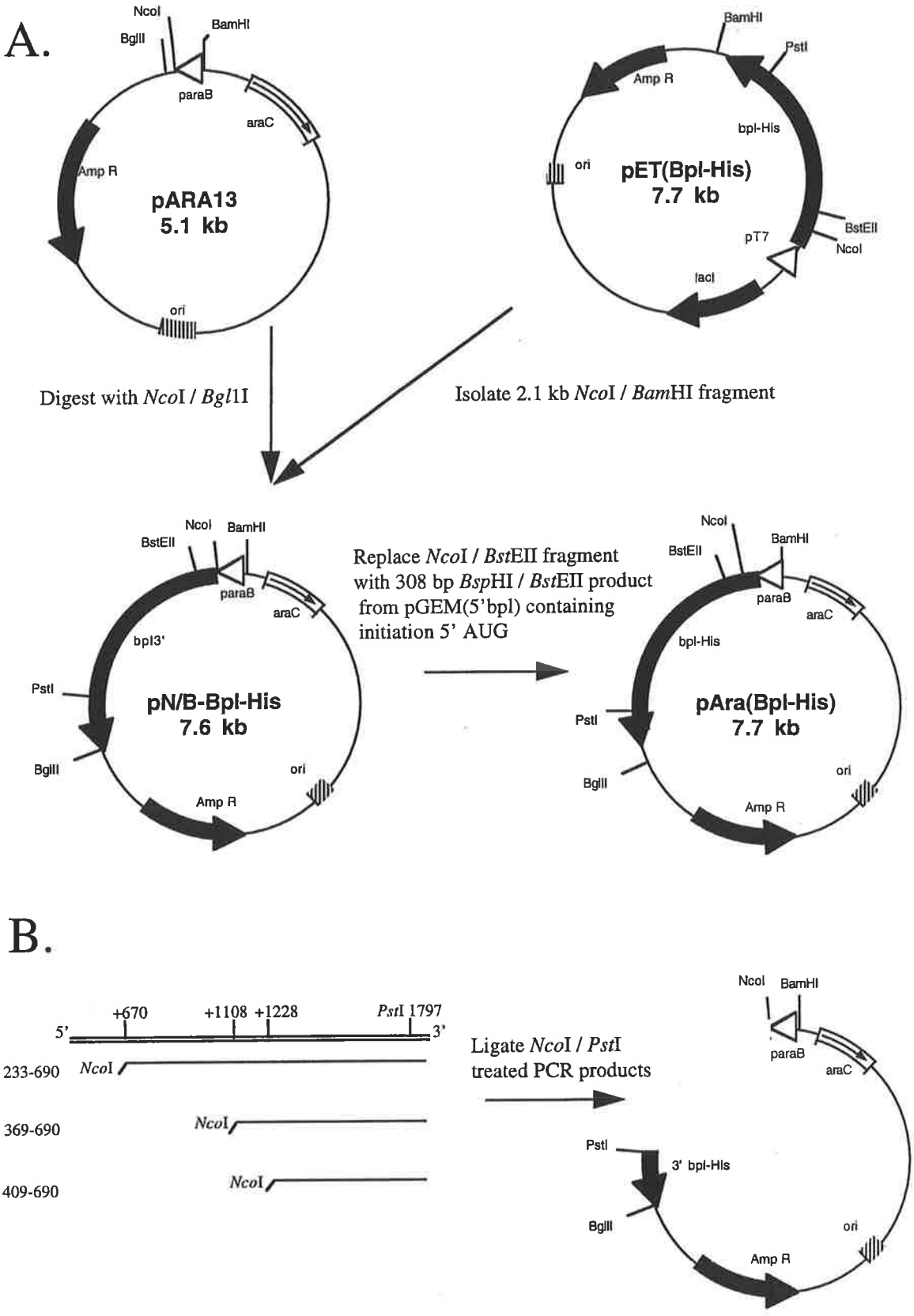

A.
NcoI
BglII
BamHI
paraB
araC
Amp R
pARA13
5.1 kb
ori
BamHI
PstI
Amp R
bpl-His
ori
pET(Bpl-His)
7.7 kb
BstEII
NcoI
pT7
lacI
Digest with NcoI / Bgl1I
Isolate 2.1 kb NcoI / BamHI fragment
NcoI
BamHI
BstEII
paraB
araC
bpl3'
PstI
pN/B-Bpl-His
7.6 kb
BglII
ori
Amp R
Replace NcoI / BstEII fragment
with 308 bp BspHI / BstEII product
from pGEM(5'bpl) containing
initiation 5' AUG
NcoI
BamHI
BstEII
paraB
araC
bpl-His
PstI
pAra(Bpl-His)
7.7 kb
BglII
ori
Amp R
B.
+670
+1108 +1228
PstI 1797
5'
3'
233-690
NcoI
369-690
NcoI
409-690
NcoI
Ligate NcoI / PstI
treated PCR products
NcoI
BamHI
paraB
araC
PstI
3' bpl-His
BglII
ori
Amp R

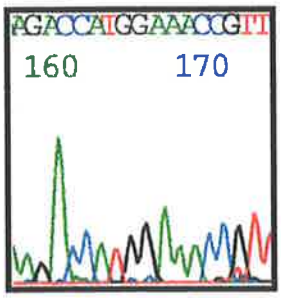
Figure 5.2: Nucleotide changes introduced to generate bacterial expression vectors for recombinant expression of N-terminally truncated variants of yBPL-His.

Mutated plasmids were sequenced using the automated dye terminator method and the chromatographs of the resolved products are shown here. The DNA sequence of the coding strand is shown above the chromatograms. The introduced *Nco*I endonuclease restriction site is underlined and the bases forming each codon are shown above the sequence. The amino acid sequence of the translated product is shown above the nucleotide sequence. Constructs A) yBPLΔ1-233, B) yBPLΔ1-369 and C) yBPLΔ1-409 are shown.

A

Met Glu[233] Thr Val

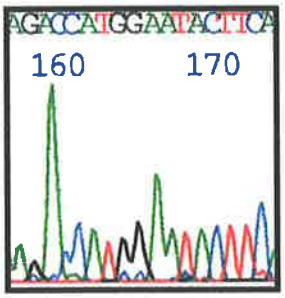
5'CCATGGAAACCGTT3'



B

Met Glu[369] Tyr Phe

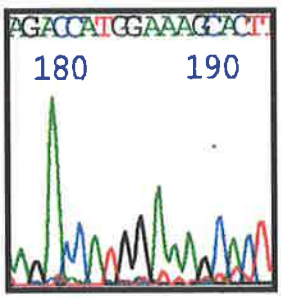
5'CCATGGAATACTTC3'



C

Met Glu[409] Ser Thr

5'CCATGGAAAGCACT3'



5.2.2 Expression and purification of yBPL truncations

The expression and purification of the yBPL truncations were performed essentially as in 3.2.2 except that the truncations were expressed in the *birA*$^+$ *E. coli* strain DH5α and induced by growth in media supplemented in 0.2% arabinose for 6 hours at 30°C. Before chromatography, cells were disrupted by two passes through a French Press (82 800 - 103 500 kPa).

5.3 RESULTS

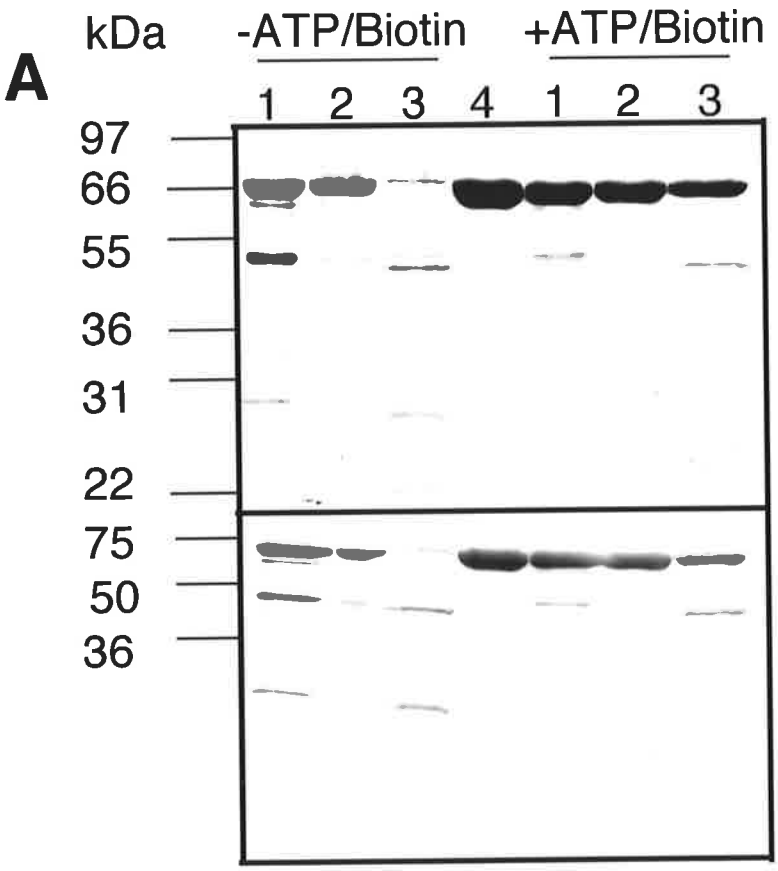
5.3.1 Limited proteolysis of yBPL

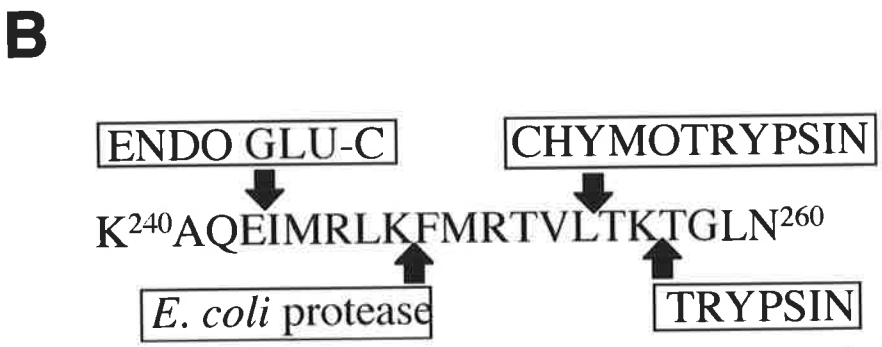
Purified intact yBPL was subjected to limited proteolysis with several proteases in order to define the domain boundaries within the enzyme. The ligase was treated with trypsin, chymotrypsin and endoproteinase Glu-C and the products analysed by SDS-PAGE (Figure 5.3A). All three proteases generated a fragment of around 50 kDa which contained the C-terminus, identified using a Ni-NTA blot to probe for the His6-tag. N-terminal sequencing of these products revealed that cleavage occurred between Lys-256 and Thr-257 for trypsin, Leu-254 and Thr-255 for chymotrypsin and Glu-243 and Ile-244 for endoproteinase Glu-C (Figure 5.3B). Digestion with both trypsin and endoproteinase Glu-C released a second fragment containing the His6-tag. These products, of approximately 27 kDa, were the result of cleavage between Pro-408 and Glu-409 with endoproteinase Glu-C and between Arg-425 and Gly-426 for trypsin. As these cleavage points are located around the predicted catalytic site, yBPL was subjected to digestion after equilibration with saturating concentrations of MgATP and biotin (Figure 5.3A). Under these conditions, the 27 kDa products were not detected, indicating that cleavage did not occur at Glu-409 or Arg-425. In addition, cleavage at Lys-256, Leu-254 and Glu-243 was considerably slower in the presence of the substrates, since the release of the 50 kDa fragment was markedly reduced in all cases. Together these data imply that the yBPL molecule contains two protease sensitive sites, one within an interdomain linker which connects a 27 kDa N-terminal domain with the remaining 50 kDa of the protein and a second region within the catalytic site. The enzyme-biotinyl-5'-AMP complex was more resistant to proteolysis at both sites than yBPL alone. This

Figure 5.3 Limited proteolysis of yBPL.

A) Yeast BPL (1.8 μM) was treated with endoproteinase Glu-C (lane 1), chymotrypsin (lane 2), trypsin (lane 3) or no protease (lane 4) at a protease to substrate ratio of 1:100 (w:w) for 2 hours at 37°C. Proteases were sequencing grade or higher (Boehringer-Mannheim). yBPL was digested either with or without equilibrating the enzyme with 3 mM MgATP and 5 μM biotin prior to addition of protease, as indicated above the lanes. Digestion was terminated by addition of SDS loading buffer and boiling for 5 minutes. Digestion products (3 μg per track) were resolved on duplicate 12% polyacrylamide gels under reducing and denaturing conditions. Total protein was visualised by Coomassie Blue R250 staining (upper), and the C-terminal His6 tag detected by Western transfer probed with Ni-NTA alkaline phosphatase (lower). The migration positions of molecular mass markers are indicated on the left.

B) The sequence of yBPL in the protease sensitive region between residues 240 and 260 is shown, with the cleavage points indicated by arrows. The position of the peptide bond cleaved by an *E. coli* protease encountered during purification, is also shown.


A
kDa
-ATP/Biotin
+ATP/Biotin
1 2 3 4 1 2 3
97
66
55
36
31
22
75
50
36


B
ENDO GLU-C
CHYMOTRYPSIN
K240AQEIMRLKFMRTVLTKTGLN260
E. coli protease
TRYPSIN

suggested that structural differences exist between the two enzyme forms, with the enzyme complex having the more compact conformation.

Tryptic digestion of either yBPL alone (apo enzyme) or yBPL-biotinyl-AMP complex (holo enzyme) was performed, measuring both the loss of the 77 kDa intact protein and enzyme activity (Figure 5.4). As expected, the holo enzyme was more resistant to proteolysis, and more enzyme activity was retained, compared to apo enzyme (Figure 5.4). For both enzyme forms, digestion of the enzyme and loss of activity occurred at slightly different rates over the first 2h whereas with longer reaction times the loss of activity corresponded to loss of intact protein. This most probably reflected the relative rates of cleavage within the two protease sensitive regions of the molecule determined above. In the initial 2h phase, cleavage occurred in the interdomain linker such that the digested enzyme retained some activity. In contrast, degradation at the secondary cleavage site within the catalytic domain inactivated the enzyme.

5.3.2 N-terminally truncated yBPL

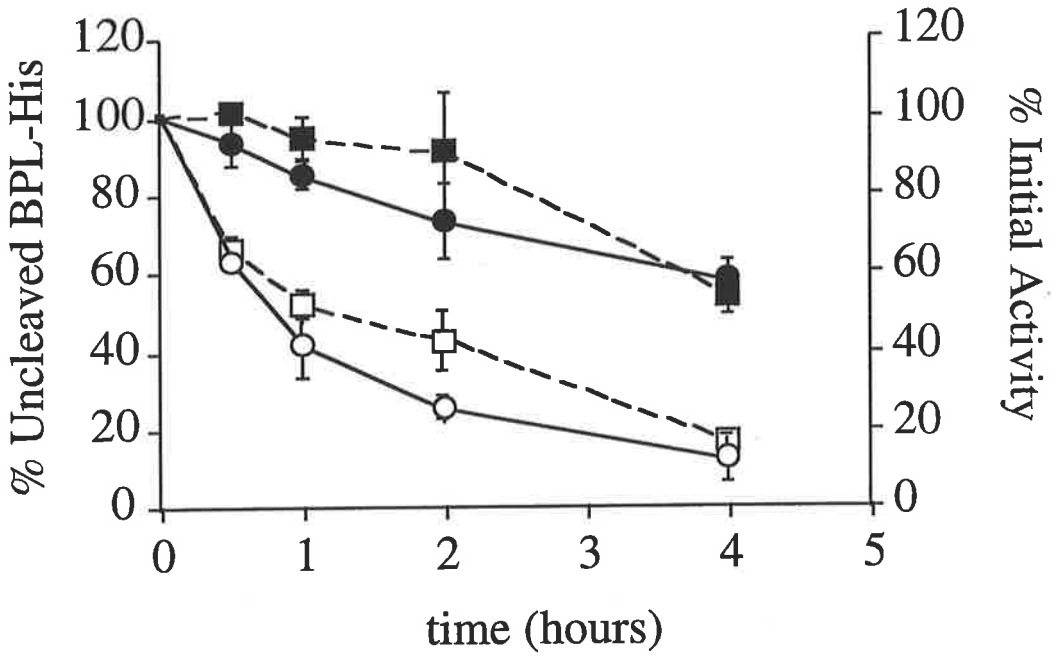
A series of N-terminally deleted forms of yBPL, all containing a C-terminal His6 tag, were designed to investigate the importance of the various regions of the molecule in catalysis. The first truncation, yBPL(Δ1-233), removed the N-terminal domain determined by domain mapping. The next truncation, yBPL(Δ1-369), contained the region of protein from residues 369 to the C-terminus, which displayed the greatest degree of amino acid homology between all BPLs (Chapman-Smith & Cronan, 1999a). Thus residue 369 lies at the N-terminal end of the proposed catalytic domain. The final truncation,

Figure 5.4. Analysis of tryptic digestion of yBPL

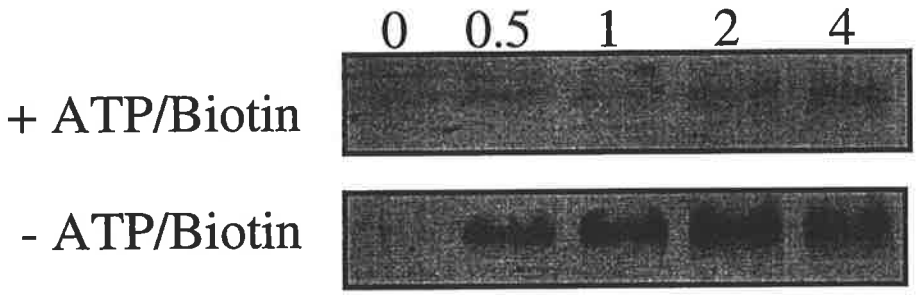
A) yBPL-His was treated with trypsin, as described in Figure 5.3, except the reactions were terminated by the addition of 0.3 μM aprotinin. The products of digestion over 4 hours were analysed by SDS-PAGE (solid line) and assayed for BPL activity (broken line). At each time point the amount of intact BPL or enzyme activity is shown as a percentage of the initial starting material. Error bars represent the standard error of at least three experiments. Prior to the addition of trypsin, yBPL was preincubated in the absence (□) or presence of (■) 3 mM MgATP and 5 μM biotin at 37°C for 5 minutes to form apo or holo enzyme respectively.

B) The formation of a 30 kDa fragment, containing the C-terminal His6-tag, was analysed by SDS-PAGE throughout the tryptic digest of yBPL both in the presence (upper) and absence (lower) of substrates. The presence of this fragment represents cleavage at the ATP-binding site in yBPL.

A


% Uncleaved BPL-His
% Initial Activity
120
100
80
60
40
20
0
0
1
2
3
4
5
time (hours)

B


0 0.5 1 2 4
+ ATP/Biotin
- ATP/Biotin

yBPL(Δ1-409), deleted part of the proposed ATP binding loop and served as a negative control. A series of vectors was produced for the expression of these truncations in *E. coli*, as described in 5.2.1 and outlined in Figure 5.1. These vectors, derived from pAra13 (Cagnon *et al.*, 1991), expressed yBPL, yBPL(Δ1-233), yBPL(Δ1-369) and yBPL(Δ1-409) under the control of an arabinose-inducible promoter.

5.3.3 *In vivo* characterisation of truncated yBPL

In order to determine the biological activity of these truncations *in vivo*, the series of expression vectors described above were transformed into the *birA,* biotin auxotroph stains strains CY918 (Chapman-Smith *et al.*, 1994) and BM4062 (Barker & Campbell, 1981b). These strains have a high requirement for biotin, and the *birA85* mutation in BM4062 also confers a temperature-sensitive phenotype. Under selective conditions only those strains expressing functional exogenous BPL survive, since the essential *E. coli* enzyme acetyl CoA carboxylase can be biotinylated and is therefore active.

The CY918 strains bearing the expression vectors were grown on permissive and selective media to test for complementation of the defective bacterial *birA1*. Strains CY918 and CY918 harbouring the parent expression vector alone, pAra13, did not grow on the selective media supplemented with 0.2% arabinose but growth was observed on non-selective media. Addition of as little as 3 nM biotin to the selective media restored growth. As expected, the strains expressing either full length yBPL or *E. coli* BirA, from plasmid pCY216 (Chapman-Smith *et al.*, 1994), permitted growth on both media. However, expression of the truncated forms of yBPL failed to complement the mutant strain

at either 30°C or 37°C on selective media (Figure 5.5A). For all truncations except yBPL(Δ1-409), an induced His6-tagged protein of the expected molecular mass was detected in crude cell lysates using Ni-NTA blot, indicating that the proteins were being expressed (Figure 5.5B). Expression of yBPL-409 was not detected possibly because this truncation, which removes part of the predicted catalytic region, is rapidly degraded.

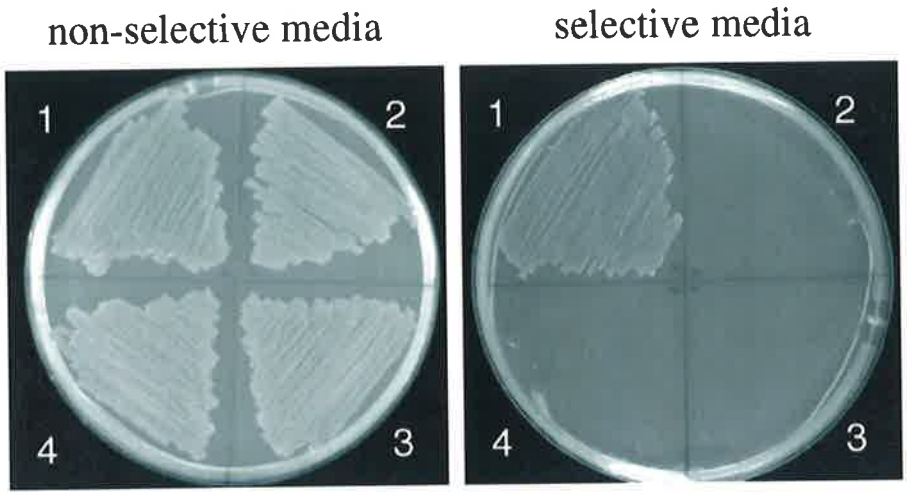
The observed failure of the yBPL truncations to complement the *birA1* mutation at low concentrations of biotin may have been due to the truncated yeast enzymes themselves having a higher biotin requirement, as has been reported for N-terminally truncated BirA (Xu & Beckett, 1996). Therefore the complementation assay was carried out in a second bacterial strain, BM4062 (Barker & Campbell, 1981b) where the endogenous bacterial BPL could be heat inactivated. At 42°C where the BirA85 protein was non-functional (Barker & Campbell, 1981b), only full length yBPL was able to sustain growth. The assay was performed on media supplemented with increasing biotin concentrations up to 1 mM. However, the additional biotin did not permit the growth of strains expressing truncated yBPL. Interestingly, the BM4062 strain expressing yBPL-369 was observed to have impaired growth on solid media even at the permissive temperature of 30°C. The reduced cellular growth of this strain was assayed by measuring the growth of the culture in liquid media at 30°C (Figure 5.6). Supplementation of 4.1 μM biotin into LB media partially alleviated the toxicity (Figure 5.6) suggesting that yBPL-369 may be sequestering free, cellular biotin. Addition of up to 41 μM biotin did not reduce the toxicity further, most probably as the bacterial biotin transport system and cellular biotin concentration were at saturation (Piffeteau & Gaudry, 1985). As the cellular biotin concentration

Figure 5.5: *In vivo* analysis of yBPL truncations in *E. coli* CY918

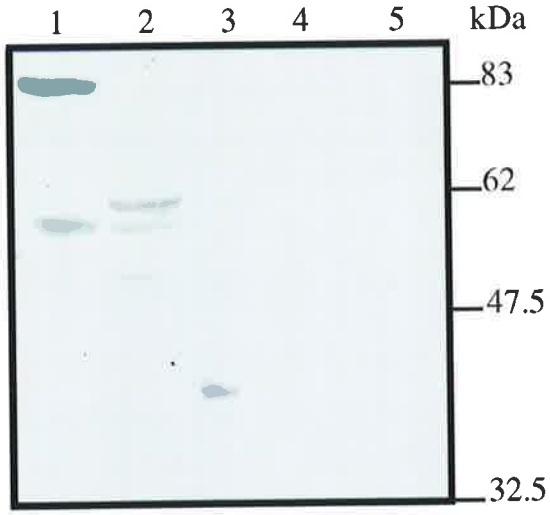
A) The *E. coli* strain CY918 was transformed with vectors for arabinose-inducible expression of N-terminal truncations of yBPL. Strains were grown on either non-selective or selective media. Strains 1 to 4 expressed full length yBPL, yBPL(Δ1-233), yBPL(Δ1-369) and yBPL(Δ1-409) respectively.

B) Whole cell lysates of *E. coli* CY918 containing the vectors for arabinose-inducible expression of yBPL truncations were fractionated on a 12% polyacrylamide gel under reducing conditions. The C-terminally His6 tagged proteins were detected by Western transfer probed with Ni-NTA alkaline phosphatase. The expressed proteins were yBPL (lane 1), yBPL(Δ1-233) (lane 2), yBPL(Δ1-369) (lane 3), yBPL(Δ1-409) (lane 4) and cells harbouring pAra13 (lane 5). The migration of molecular mass markers is indicated on the right.

A


non-selective media
selective media
1
2
4
3
1
2
4
3

B


1
2
3
4
5
kDa
83
62
47.5
32.5

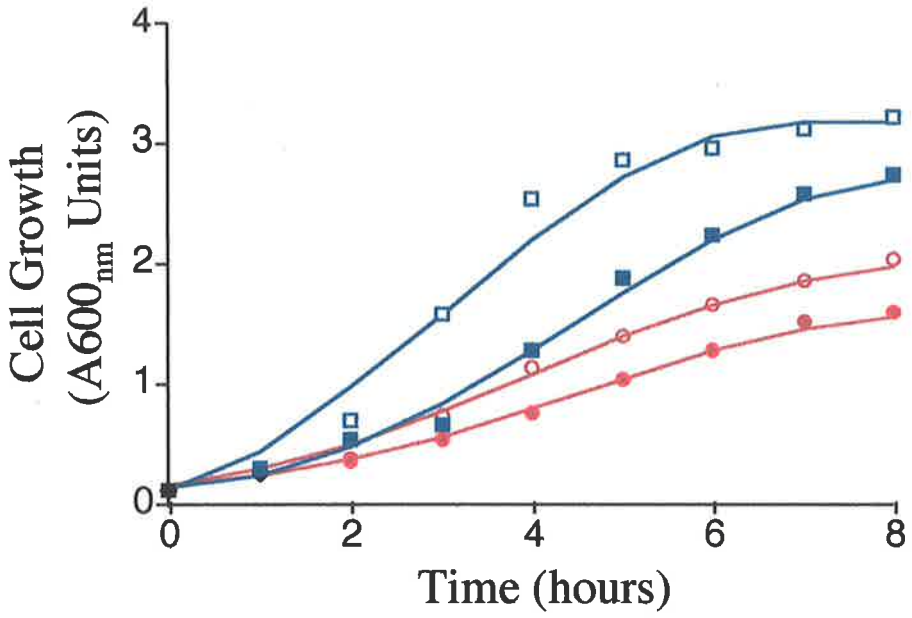


Figure 5.6: Toxicity of yBPL(Δ1-369)His6

BM4062 cells harbouring the plasmids indicated were cultured at 30°C in LB media containing 100 μg/ml ampicillin and 0.2% arabinose and growth was monitored at 600nm. Cells expressing yBPL(D1-369) (red lines) were grown in media alone (●) or media supplemented with either 4.1 μM or 41 μM (○, identical growth curves). Similarly cells harbouring pAra13 (blue lines) were grown in media alone (■) or media supplemented with 4.1 μM biotin (□).

of the host cells could not be increased for the *in vivo* analysis, an alternative approach was to increase the cellular concentration of BCCP. It was reasoned that by increasing the concentration of BCCP, there would be more opportunity for the cell to produce active acetyl CoA carboxylase if yBPL-369 biotinylated the protein substrate. Co-expression of the BCCP and biotin carboxylation subunits of acetyl CoA carboxylase from the plasmid pLS21 (Li & Cronan, 1992) with yBPL-369 also did not improve cellular growth. Together these data imply that yBPL-369 may behave as a biotin binding protein that is compromised in BPL activity.

5.3.4 *In vitro* characterisation of truncated yBPL

The yBPL truncations were expressed in *E. coli* DH5α cells and partially purified by nickel chelating chromatography. The Ni-NTA purified material recovered from cells harbouring pAra13 displayed no BPL activity, showing that endogenous bacterial BPL had been removed. Full length yBPL was found to have the highest specific activity (25 nmol/min/mg) (Figure 5.7). Removal of the N-terminal domain in construct yBPL(Δ1-233) was found to reduce the activity of the enzyme by greater than 3500-fold (7 pmol/min/mg) (Figure 5.7). The truncation yBPL(Δ1-369) had very low specific activity (1.7 pmol/min/mg). As expected yBPL(Δ1-409), which has part of the proposed ATP binding motif deleted, showed no activity. It is evident from the *in vivo* complementation assays that the 3500-fold reduction in activity seen on the deletion of the N-terminal domain reduced the activity of yBPL to a level that is inadequate for viability.

The yBPL(Δ1-233) truncation was further purified by anion exchange chromatography and tested for activity in the presence of increasing concentrations of biotin. As the addition of increasing amounts of labelled [^{3}H]biotin was found to contain some inhibitory contaminant, BPL assays with large quantities of cold biotin were performed by analysis of holo BCCP formation using native non-denaturing PAGE. Biotinylation of the molecule results in the formation of neutral charge of the biotin-accepting lysine side chain allowing the separation of apo and holo protein based on this single charge difference. The addition of up to 500 μM of biotin in the reaction did not improve the enzyme's activity (Figure 5.8) suggesting the observed decrease in activity for this truncation was not the result of an increased K_m for biotin, consistent with the *in vivo* complementation assays.

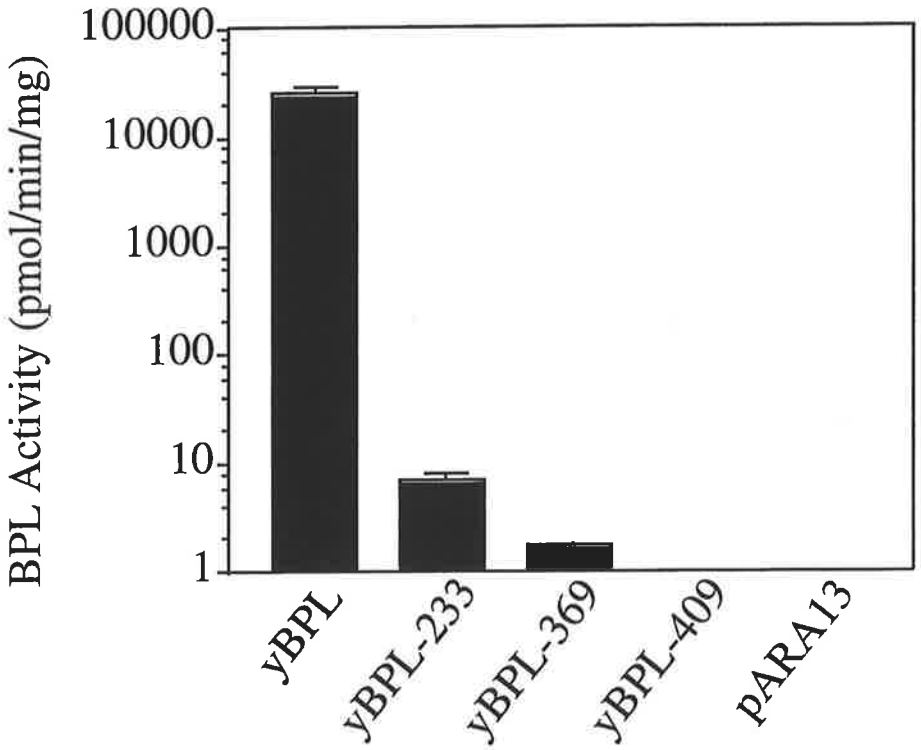


Figure 5.7: *In vitro* assays of truncated forms of yBPL

The N-terminally truncated forms of yBPL were expressed in *E. coli* DH5α and the protein recovered after NiNTA chromatography was assayed in the *in vitro* biotinylation reaction described in 3.2.3

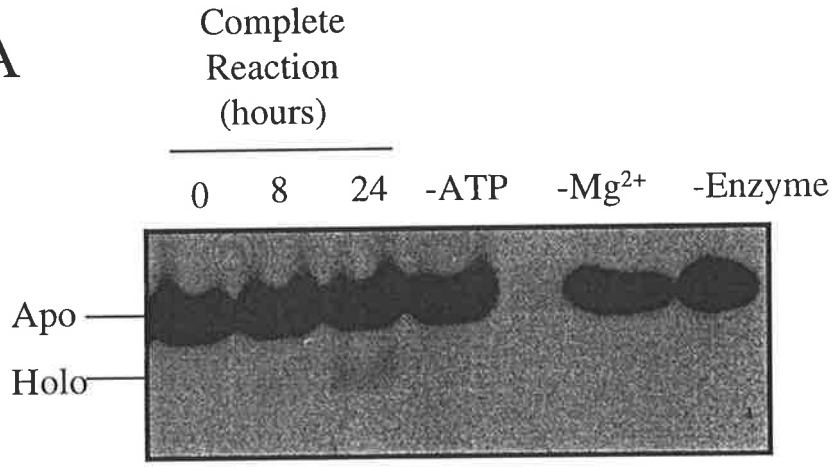
Figure 5.8 Activity of yBPL(Δ1-233).

Biotinylation by yBPL(Δ1-233) was assayed *in vitro* using apo BCCP-87 as the biotin accepting substrate. The products of the reaction were fractionated by non-denaturing PAGE using 15% polyacrylamide gels and Tris-tricine running buffer (Schagger & von Jagow, 1987; Chapman-Smith *et al.*, 1994) and the proteins detected by straining with Coomassie blue R250. To the left of the gels, the migratory position of apo and holo BCCP-87 are shown.

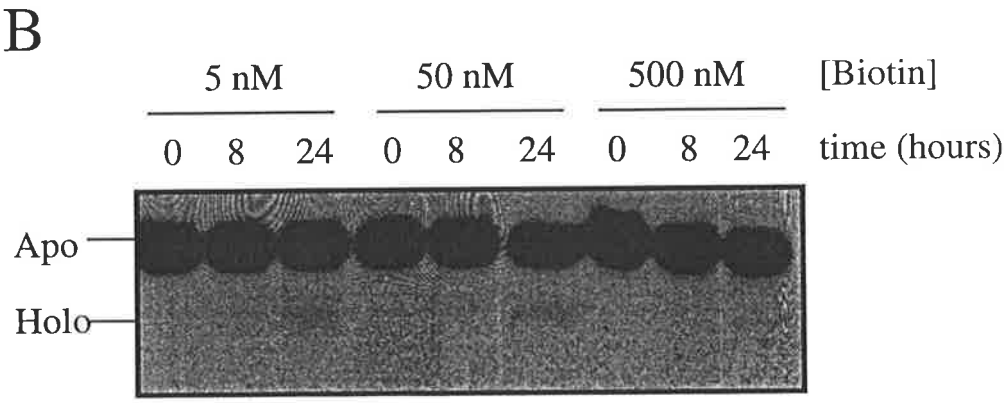
A) The generation of holo BCCP-87 over time with the complete assay mix is shown, as is the absence of products after 24 hours when ATP, Mg^{2+} and enzyme were omitted.

B) The activity of yBPL(Δ1-233) with increasing quantities of biotin, indicated above the time point values, is shown

A


Complete Reaction (hours)
0
8
24
-ATP
-Mg2+
-Enzyme
Apo
Holo

B


5 nM
50 nM
500 nM
[Biotin]
0 8 24 0 8 24 0 8 24
time (hours)
Apo
Holo

5.4 DISCUSSION

Limited proteolysis of purified yBPL, undertaken to determine the domain structure of the protein, indicated that the protein contained two protease sensitive regions, the more N-terminal of which lies between residues 240 and 260. The observation that all three proteases tested, as well as an endogenous *E. coli* protease, cut the protein within this sequence suggests that this region most probably forms an exposed linker sequence between a 27 kDa N-terminal domain and the remaining 50 kDa portion of the protein. The second protease sensitive region observed in yBPL occurred within the 50 kDa C-terminal region which, based on sequence homologies, contains the catalytic centre (Cronan & Wallace, 1995). In fact, both the trypsin and endoproteinase Glu-C sensitive residues mapped in the putative ATP and biotin binding region of the catalytic domain, with the trypsin cleavage site within the proposed ATP binding motif, GRGRGG (Cronan & Wallace, 1995). These cleavage sites were protected from proteolysis and subsequent loss of yBPL activity in the presence of ATP and biotin (Figures 5.5 and 5.6). Correspondingly, the ligand-bound form of yBPL was more resistant to the loss of enzyme activity which accompanied proteolysis than the apo enzyme. These observations support the identification of the catalytic site made from sequence homologies. In the crystal structure of BirA, the region containing the ATP binding motif is one of several poorly defined, solvent exposed loops found close to the catalytic centre (Wilson *et al.*, 1992). A subtilisin-sensitive site lies within an unstructured loop adjacent to the ATP binding loop, and cleavage is inhibited when either biotin or biotinyl-5'-AMP is bound to the enzyme (Xu *et al.*, 1995). The results of limited proteolysis presented here indicate that yBPL forms two domains, and that the cleavage

observed around the ATP binding motif suggests the presence of an exposed loop structure within the 50 kDa C-terminal domain, as is seen in BirA.

In addition, and somewhat surprisingly, the presence of ATP and biotin also reduced protease susceptibility within the proposed linker sequence between the two domains. This suggests that substrate binding caused a global conformational change, affecting sequences at some distance from the binding site in the primary structure. Conformational changes in BirA associated with substrate binding have also been demonstrated. The exposure of crystals of BirA to biotin caused cracking of the crystals (Wilson *et al.*, 1992). Interestingly, the N-terminal domain of BirA which is involved in DNA binding appears to affect the affinity of the catalytic domain for both biotin and biotinyl-5'-AMP. A truncation mutant in which the N-terminal domain was absent was still able to catalyse biotin transfer, but displayed a 100-fold decrease in the affinity for biotin and a 1000-fold decrease in the affinity for biotinyl-5'-AMP (Xu & Beckett, 1996). Biotin binding to this truncation caused no quenching of intrinsic protein fluorescence, as opposed to the 15% quenching observed with the intact enzyme (Xu & Beckett, 1996). These data suggest that quenching of fluorescence may be the result of the conformational changes which are induced by biotin binding and the truncated enzyme is compromised in its ability to go through these changes. The interaction between the DNA binding domain and the catalytic domain is accompanied by conformational changes and is thought to relate to repressor function of BirA (Wilson *et al.*, 1992). The data presented here indicate that although the N-terminal domain in yBPL has no equivalent DNA binding function, there is a functional interaction with the catalytic domain.

Furthermore, expression of N-terminally truncated variants of yBPL in two *E. coli* strains carrying *birA* mutations showed that the presence of both domains

was necessary to produce a functional enzyme. The analysis of the activity of the yBPL truncations *in vitro* is consistent with the results of the complementation assays, and it is evident that the 3500-fold reduction in activity seen on the deletion of the N-terminal domain reduced the activity of yBPL to a level that is inadequate for viability. It is likely that in the absence of the N-terminal domain the conformational changes associated with substrate binding, and apparently necessary for enzymatic activity, may occur at a slower rate than in the presence of the domain and therefore effect the overall activity of the protein. This agrees with the observation that tryptic cleavage of yBPL in the linker region produced a form of yBPL which retained some biological activity in *in vitro* assays (Figure 5.6). While the precise role of the N-terminal domain is unclear, the results presented here are consistent with the studies of known defects in human BPL. Several point mutations in the N-terminal domain of human BPL result in a defective enzyme (Suzuki *et al.*, 1994; Dupuis *et al.*, 1996; Aoki *et al.*, 1997), indicating that the integrity of this region of the protein is important for function. Sequence homology between yeast and human BPL in the N-terminal domain is low and allows different alignments, making it difficult to precisely identify analogous residues and therefore to produce point mutations in yBPL which mimic those isolated in the defective human enzyme. However, these mutations in human BPL are found upstream of the protease sensitive linker region in yBPL identified here (Chapman-Smith & Cronan, 1999a), consistent with the inability of the N-terminal truncation yBPL(Δ1-233) to complement the *birA* defects. Structural characterisation of the N-terminal domain and the identification of interactions with both the catalytic site and other molecules will aid in determining the role of this domain in enzyme function.

Chapter 6

ESTABLISHING A NOVEL SELECTION TECHNIQUE FOR INVESTIGATING *IN VIVO* BIOTINYLATION

ESTABLISHING A NOVEL SELECTION TECHNIQUE FOR INVESTIGATING *IN VIVO* BIOTINYLATION

6.1 INTRODUCTION

A novel, genetic selection for the isolation of BCCP-87 mutants with a decreased affinity for BirA was established in the laboratory of Dr John Cronan Jnr. at the Department of Microbiology and Biochemistry, University of Illinois. Our laboratory is involved in a collaboration to analyse the properties of these mutant peptides and current work by Dr Anne Chapman-Smith is investigating the determinants on both BCCP-87 and BirA necessary for governing substrate specificity. Given the success of the *in vivo* selection and our laboratory's on-going interest in understanding the structure and function of pyruvate carboxylase, we decided to investigate biotinylation using the yeast PC biotin domain.

In the genetic selection reported by Chapman-Smith *et al.* (1999), several of the mutations induced gross conformational changes to the otherwise structured biotin domain. This rendered the peptide a low affinity substrate for BirA. One mutant in particular, G133S, was a known temperature sensitive substrate, poorly biotinylated at 37°C (Li & Cronan, 1992). In order to remove these gross structural mutants from a mutagenic library, phage display was incorporated into the selection procedure. Phage display would permit the isolation of mutants able to interact with BirA by selecting biotinylated phagemids, the product of enzymatic biotinylation. This pool of functional mutants would then be subjected to the genetic selection to specifically isolate low affinity

substrates. This chapter describes the work performed using the yeast PC biotin domain (yPC-104) to establish phage display and the *in vivo* selection.

6.2 SPECIFIC METHODS

6.2.1 DNA manipulations and sequencing

Polymerase chain reactions, using 100 ng of oligonucleotides with 10 ng of template, were performed in 40µl of 1x thermophilic buffer (ProMega) containing 2.5 - 4 mM $MgCl_2$, 5 mM dNTPs and 1U *Taq* DNA polymerase with thirty cycles of denaturation 95°C 1 min ; annealing 60-65°C 1.5 min ; and synthesis 72°C 1.5 min. Sequencing of all the constructs was performed using either the Sequenase Version II Kit (Pharmacia) or ABI Prism Dye Terminator sequencing (Perkin Elmer).

6.2.1.1 Construction of phage display vector

DNA encoding yPC-104 was amplified using PCR with the oligonucleotides YPC104B and YPC3'B, with template pMW4A which contains the cDNA for *S. cerevisiae* pyruvate carboxylase 1 (Lim *et al.*, 1988). The primers introduced *Nco*I and *Pst*I endonuclease restriction sites respectively, facilitating the cloning of the 330 bp product into the phagemid pGF-14 (Lucic *et al.*, 1998), a derivative of pHEN-1 (Hoogenboom *et al.*, 1991). This generated the vector pFdBD-104. The biotinylated Lys residue at position 1135 in pyruvate carboxylase 1 was replaced with a Leu residue in the construct pFdBD-K1135L. The mutant was obtained by PCR using the oligonucleotides Lys61Leu and YPC3'B and pFdBd-104 as the template. The mutated DNA was digested with endonuclease restriction enzymes *Bcl*I and *Pst*I and ligated into similarly treated pFdBD-104. Constructs pFdBD-104 and pFdBD-K1135L were transformed into *E. coli* XL1 Blue (Stratagene) for phage display.

6.2.1.2 Construction of bacterial expression vector

DNA encoding yPC-104 was cloned into the expression vector pKK223-3 (Pharmacia). Oligonucleotide YPC104C was employed, using PCR with YPC3'B and pFdBd-104, to linker an *Eco*RI endonuclease restriction site onto the 5' end of DNA encoding yPC-104. The product was digested with *Eco*RI and *Pst*I and cloned into similarly treated pKK223-3, generating the construct pC-104. This vector was subsequently digested with *Nco*I and *Pst*I and ligated to the 330 bp *Nco*I / *Pst*I product liberated from pFdBd-K1135L, yielding the construct pC-K1135L. Proteins expressed from these vectors had an additional Met-Ala motif on the N-terminus, the result of cloning extra nucleotides at the 5' region of the gene introducing the *Nco*I restriction site.

6.2.2 Expression of gIIIp fusion proteins

XL1 Blue cells harbouring either pFdBD-104 or pFdBD-K1135L were grown in 2YT supplemented with ampicillin (100 µg/ml), tetracycline (10 µg/ml) and 10 µM biotin. Overnight cultures were diluted 1:50 in fresh media, grown to log phase at 37°C and fusion protein expression induced with 0.02 mM IPTG for 16 hours at 30°C. Cell lysates were prepared as described by Chapman-Smith *et al.* (1994) and analysed by SDS-PAGE on 12% polyacrylamide gels (Laemmli, 1970). Biotinylated proteins were detected with Western blotting as described by Lim *et al.* (1987) and gIIIp fusion proteins detected using anti gIIIp antibody (Mo Bi Tec).

6.2.3 Preparation of phagemids

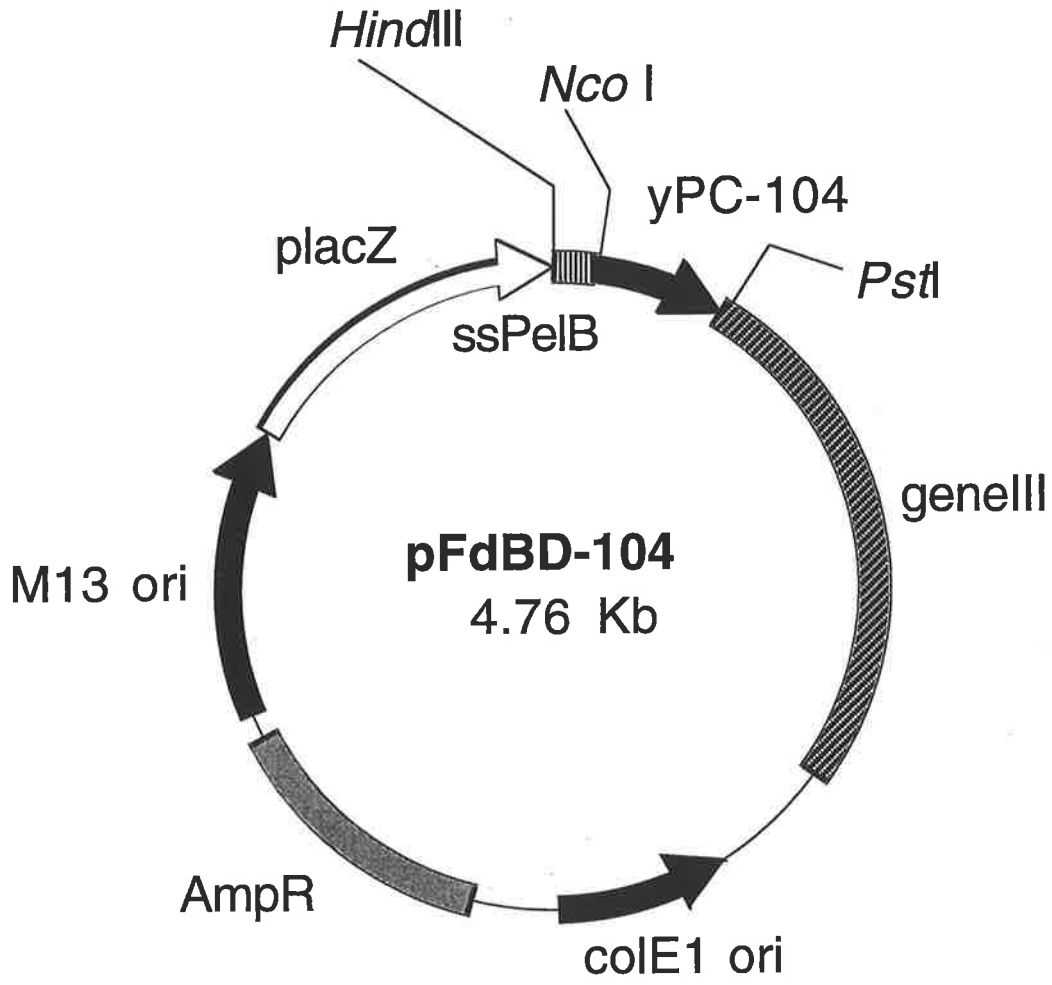
Overnight cultures of XL1 Blue cells harbouring phagemid vectors were grown in 2YT supplemented with 100µg/ml ampicillin, 10 µg/ml tetracycline, 10 µM biotin and 2% glucose, were diluted 1:50 in fresh media and grown at 37°C to A_{600nm} of 1.2. The cells were subcultured, 200 µl into 20 ml of fresh media from which glucose was omitted, together with 6 x 10^9 plaque-forming units of M13K07 helper phage. At A_{600nm} of 0.8 the culture was supplemented with 0.02 mM IPTG and kanamycin (50 µg/ml) and shifted to 30°C for 16 hours. Phagemids were prepared from cleared culture supernatant using PEG precipitation (Lowman & Wells, 1991).

6.2.4 Biopanning

Phagemids were panned using magnetic beads coated with Streptavidin (Dynal). For diagnostic panning, 30 µg of beads were incubated with 10^8 phagemids in 1 ml PBS / 0.1% BSA for 1 hour at room temperature, captured using a magnet and washed 10 times in 1 ml wash buffer (PBS / 0.1% BSA / 0.5% Tween-20). After washing, beads were resuspended in PBS / 0.1 % BSA and the bead suspension added to log phase XL1 Blue cells for 15 minutes at 37°C. Bacteria infected with phagemids were detected by growth on LB media supplemented with 100 µg/ml ampicillin.

Figure 6.1 Plasmid Map of pFdBD-104.

The plasmid pFdBD-104 was constructed from pGF14 (Lucic *et al.*, 1998), a derivative of pHEN (Hoogenboom *et al.*, 1991). DNA encoding the biotin domain of yeast pyruvate carboxylase 1 (yPC-104) was fused at the 5' end with the DNA encoding the secretion signal of PelB (ssPelB). An amber stop codon and the codons for a H64A subtilisin BPN' cleavage motif separate the 3' end of filamentous bacteriophage gene III. Expression of the ssPelB-yPC104-gIIIp fusion protein is under the control of the lacZ promoter (placZ). A bacterial origin of replication (colE1) permits replication of the plasmid in *E. coli* and a bacteriophage origin (M13 ori) facilitates the production of single stranded DNA for packaging into phagemids. An ampicillin resistance gene (AmpR) allows selection of bacteria harbouring pFdBD-104. The restriction sites used for constructing the plasmid are shown.


HindIII
Nco I
yPC-104
placZ
PstI
ssPelB
geneIII
pFdBD-104
4.76 Kb
M13 ori
AmpR
colE1 ori

6.3 RESULTS AND DISCUSSION

6.3.1 Construction of phage display vectors

The construct pFdDB-104 was generated, as described in 6.2.1.1 and shown in Figure 6.1, for the expression of yPC-104 as a fusion to the gene III coat protein (gIIIp) of fd filamentous bacteriophage. An amber stop codon immediately after the DNA encoding yPC-104 but preceding gene III permitted fusion of the biotin domain onto the N-terminus of gIIIp in *supE* host strains (Hoogenboom *et al.*, 1991). The two moieties were fused via a flexible glycine rich linker containing a H64A subtilisin BPN' protease sensitivity motif (Polyak *et al.*, 1997). The fusion also contained a periplasmic signal sequence necessary for phage assembly (Makowski, 1994). When co-expressed in *E. coli* with viral proteins supplied by helper phage, the fusion protein should be packaged into phagemid particles with yPC-104 displayed on the virion surface. A second construct, pFdBD-K1135L, in which the codon for the biotinylated Lys was substituted with a codon for Leu, but otherwise identical to pFdBD-104, was produced to serve as a negative control in biopanning procedures.

6.3.2 Expression and biotinylation of gIIIp fusion proteins

To investigate the expression and *in vivo* biotinylation of gIIIp fusion proteins, XL1 Blue cells harbouring either pFdBD-104 or pFdBD-K1135L were treated with 20 μM IPTG for 16 hours to induce expression of the fusion protein. Whole cell lysates were analysed by SDS-PAGE and Western transfer, probing with either anti gIIIp antibody or avidin-alkaline phosphatase conjugate to detect covalently attached biotin (Figure 6.2). As expected from previous work (Val *et*

al., 1995), yPC-104 produced from the phage display vector was biotinylated by BirA *in vivo*. The antibody blot (Figure 6.2, panel C) detected expression of gIIIp as fusions to both yPC-104 and Lys1135→Leu. The deduced molecular weight of the fusion proteins (87 kDa) was anomalously high compared with the expected molecular weights for the proteins (53.2 kDa) but is consistent with various reports which show that in SDS-PAGE analysis, gIIIp migrates to an apparent molecular weight larger than expected (Gray *et al.*, 1981; Crissman & Smith, 1984). Proteolysis of both fusion proteins was evident, suggesting the gIIIp moiety was susceptible to degradation by *E. coli* enzymes *in vivo* (Lowman, 1997; Winter *et al.*, 1994). Biotin containing bands comigrating with gIIIp bands were present only from cells harbouring the construct pFdBD-104 (Figure 6.2, panel B, lane 2) with the strongest signal from one of the presumed proteolytic products. Because it had been previously reported that biotinylated yPC-104 is relatively protease resistant (Val *et al.*, 1995), this suggested that degradation of the fusion protein occurred within the C-terminal gIIIp moiety. An additional biotin containing band was detected around the expected molecular weight for the non-fused biotin domain (11.4 kDa), the result of termination of translation in the *supE* host. In addition endogenous BCCP was detected (21 kDa) but resolved poorly away from yPC-104 under the conditions used for SDS-PAGE. These results show that when fused N-terminally to gIIIp, yPC-104 was expressed in a conformation that was recognised by BirA. Moreover, analysis of periplasmic extracts detected biotinylated yPC-104 in XL1 Blue cells grown at 30°C. As expected, substitution of the modified Lys residue with Leu abolished biotinylation, Thus the Lys1135→Leu isolated domain and gIIIp fusion protein were not detected with avidin-alkaline phosphatase (Figure 6.2, panel C, lane 3).

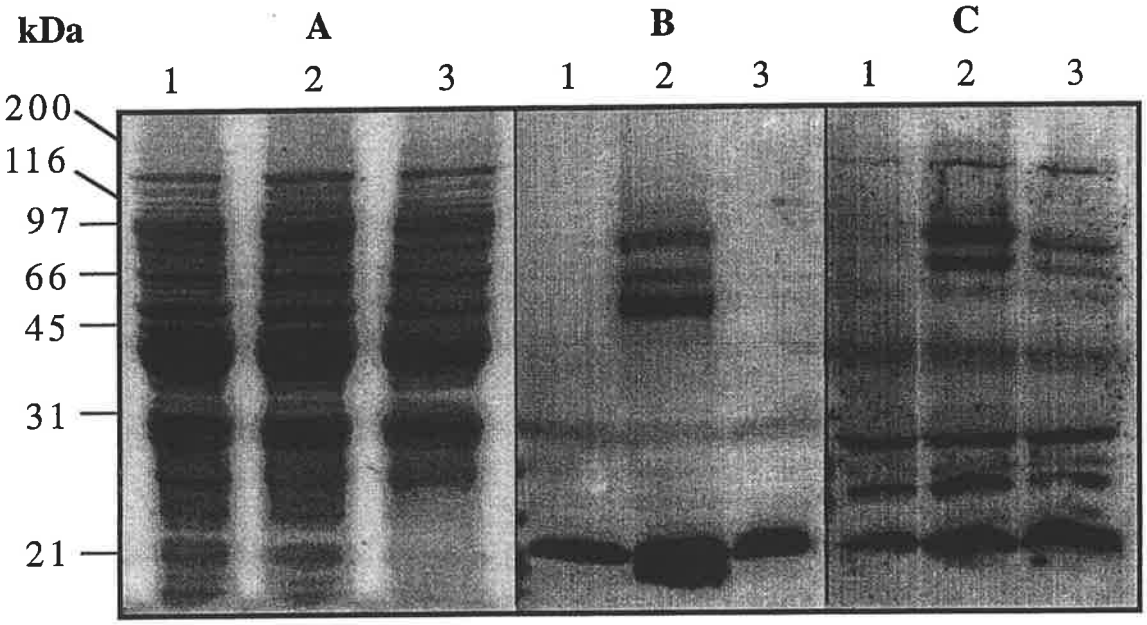


Figure 6.2: Expression of gIIIp fusion proteins

Crude cell lysates containing gIIIp fusion proteins were prepared and fractionated on triplicate 12% polyacrylamide gels under reducing, denaturing conditions. The figure shows crude lysates from *E. coli* XL-1Blue cells harbouring (1) no plasmid, (2) pFdBD-104 and pFdDB-K1135L. Proteins were detected by (A) Coomassie brilliant Blue R250 staining, (B) avidin-alkaline phosphatase and (C) anti gIIIp antibodies. The migration of molecular mass markers is indicated on the left.

6.3.3 Display of the biotin domains on phagemids

By expressing gIIIp fusion proteins at low levels in the presence of M13K07 helper phage, monovalent display of proteins on filamentous phage is possible (Hoogenboom *et al.*, 1991). Phagemids displaying the yPC-104 fusion protein were prepared under conditions known to allow monovalent display, captured using magnetic beads coated with streptavidin and used to infect male (F^+) *E. coli* cells as described in 6.2.4. Because of the high affinity of the biotin-streptavidin interaction (Wilchek & Bayer, 1990), conventional elution of captured phagemids was not possible. Instead, phagemids coupled on the beads were mixed directly with *E. coli* cells. Bacteria infected in this way were transduced to ampicillin resistance and assayed by growth in the presence of ampicillin on LB media. To test the efficiency of this panning procedure, yPC-104 and Lys1135→Leu phagemids were analysed against a null phagemid, namely Bluescript (Table 6.1). The results show that a single panning step produced a 5000 fold enrichment of the biotinylated yPC-104 phagemids over the null control. Identical low recoveries of phagemids displaying the non-biotinylated mutant and the null control were observed. Thus the interaction of the phagemids with the streptavidin coated beads, and their subsequent capture, was specific for the presence of a biotin group covalently attached to the phagemids via the Lys found at position 1135 of pyruvate carboxylase 1.

Table 6.1 : Selection of phagemids displaying biotin group

	Bluescript	yPC-104	Lys1135→Leu
In	1×10^8	1×10^8	1×10^8
Out	<10	5×10^4	<10
% Recovery	<0.00001	0.05	<0.00001

yPC-104, Lys1135→Leu or Bluescript phagemids (1×10^8 colony forming units) were added to streptavidin coated magnetic beads and panned as described in 6.2.4. The number of phagemids included in (In) and obtained from (Out) the panning experiments are shown. The results shown are the mean of three independent experiments where samples were analysed in triplicate.

6.3.4 Establishing the *in vivo* selection

A selection procedure for the isolation of mutants of the biotin domain of *E. coli* BCCP which are defective for biotinylation has been previously established (Chapman-Smith *et al.*, 1999). Briefly, mutants of the biotin domain of BCCP were overexpressed in the biotin auxotroph *E. coli* strain TM21 and grown on media containing limited biotin (4.1 nM). High affinity substrates for biotin ligase were toxic to the host as these peptides reduced the cellular pool of free biotin by converting it into a non-available, protein-bound form. The isolation of low affinity BirA substrates was achieved as mutants which exhibited a decrease in affinity for the enzyme relieve the selective pressure applied to these cells. The result was biotinylation of essential acetyl CoA carboxylase and cell growth. In order to use this approach to isolate mutants from the phage display library it was first

necessary to establish that wildtype yPC-104 competed efficiently with endogenous BCCP for biotinylation.

Initially the phage display constructs pFdBD-104 and pFdBD-K1135L were transformed into TM21 for expression of the wildtype and non-biotinylated proteins. From these vectors the secreted forms of yPC-104 and Lys1135→Leu respectively were produced without the gIIIp fusion. The bacterial strains were grown on selective and non-selective media and the growth phenotypes observed (Figure 6.3). Under the growth conditions used, expression of either yPC-104 or Lys1135→Leu did not inhibit growth. Inhibiting the secretion machinery by supplementing the media with sodium azide (Oliver *et al.*, 1990) was also ineffective in inhibiting growth. In order for the *in vivo* selection to function, it is necessary for the recombinantly expressed biotin domains to titrate the available pool of free biotin supplied to the bacterial strain. To achieve this with our system, the yeast proteins required higher level expression and increased time of exposure to cytoplasmic bacterial biotin ligase. Therefore DNA encoding the biotin domains was introduced into a second vector, pKK223-3, which contains the strong, IPTG inducible *tac* promoter (Brosius & Holy, 1984) and lacks a secretion signal. This yielded the constructs pC-104 and pC-K1135L. Expression of yPC-104 from pC-104 was found to be lethal to TM21 cells grown on limited biotin media supplemented with 2 mM IPTG (Figure 6.3). Restricted growth was observed on either limited biotin media or non-selective media supplemented with IPTG but both conditions were required for growth inhibition. The non-biotinylated Lys1135→Leu, expressed from pC-K1135L, did not affect growth on any of these media. This indicated that the biotin domain of yeast pyruvate carboxylase, expressed from the pKK223-3 derived vectors, competed

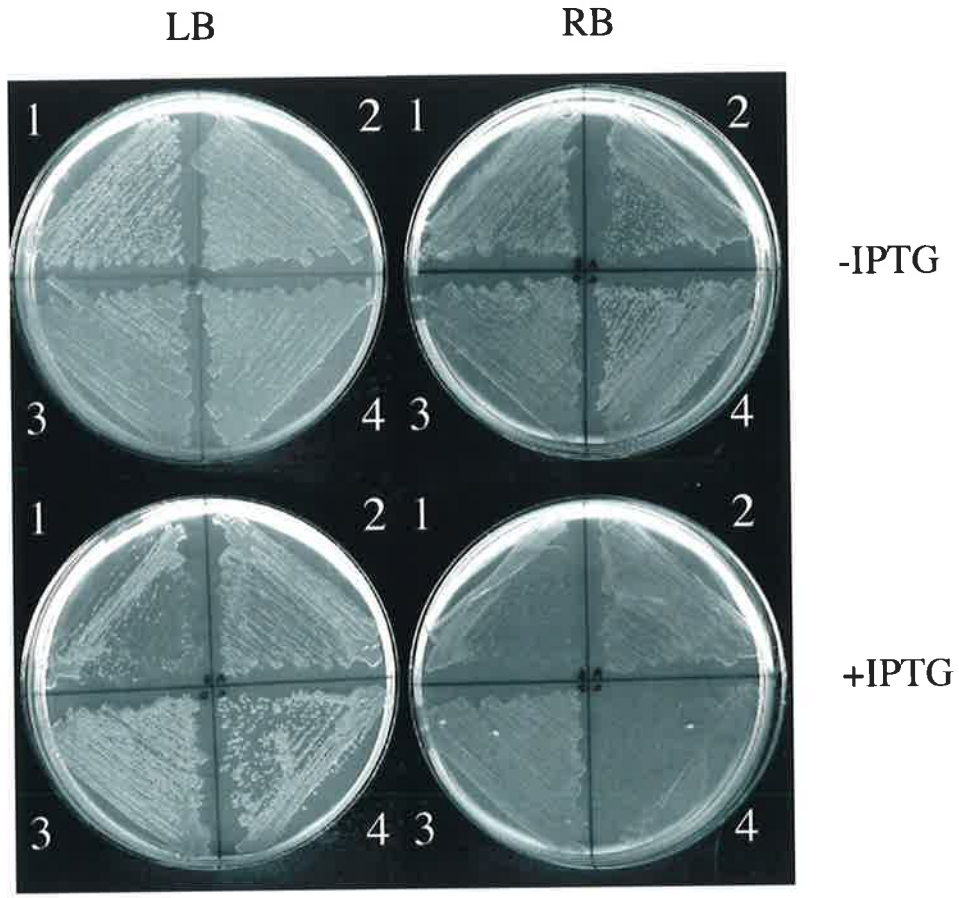


Figure 6.3: Establishment of an *in vivo* selection in *E. coli* to isolate yPC-104 mutants

An *in vivo* selection, using biotin auxotroph *E. coli* as hosts to screen for biotin domain mutants with lowered affinity for BirA, has been described previously (Chapman-Smith *et. al*, 1999) TM21 cells harbouring the plasmids (1) pFdBD-K1135L, (2) pFdBD-104, (3) pC-K1135L or (4) pC-104 were grown on either LB (non-selective media) or RB (selective media containing limited biotin). The strains were grown either without or with of 2 mM IPTG added to the media.

effectively with endogenous BCCP for free biotin and thus could be used in the *in vivo* selection.

6.3.5 Conclusions

The biotin domain of yeast PC1 can be expressed and biotinylated *in vivo* by bacterial BirA when fused to a coat protein of filamentous bacteriophage. This post-translationally modified fusion protein was successfully packaged into phagemid particles upon infection of the bacterial cells with helper phage. By employing streptavidin-coated magnetic beads, it was demonstrated that phagemids containing a biotin moiety could be selectively isolated from phagemids displaying either a non-biotinylated mutant biotin domain or no foreign protein. However, the phage display vector was not suitable for the expression of yPC-104 in the *in vivo* selection. A higher level expression vector was constructed with convenient restriction sites permitting subcloning of DNA encoding yPC-104 from the phage display vector. Expression of the wildtype peptide from this plasmid inhibited the growth of the host bacterial strain grown on biotin starved conditions whereas the non-biotinylated protein did not. Chapter 7 describes the mutational analysis of yPC-104 by employing the phage display and genetic selection systems established here.

Chapter 7

MUTATIONAL ANALYSIS OF THE YEAST PC1 BIOTIN DOMAIN

MUTATIONAL ANALYSIS OF THE YEAST PC1 BIOTIN DOMAIN

7.1 INTRODUCTION

In Chapter 6, the techniques of phage display and an *in vivo* selection in *E. coli* were established for the biotin domain of yeast PC1. In this chapter, these techniques were coupled allowing a mutational analysis of yPC-104 to specifically isolate peptides with low affinity for biotin ligase. A library of randomly mutated peptides was created and subjected to phage display followed by the genetic selection. Interesting peptides were analysed further by *in vivo* and *in vitro* biotinylation assays.

A molecular model of yPC-104 was performed in collaboration with Dr Terry Mulhern, Department of Biochemistry, University of Adelaide. A description of this model is included in this chapter as it provided a valuable tool in interpretating the results obtained from the mutational analysis.

7.2 SPECIFIC METHODS

7.2.1 Construction of mutagenic library

DNA encoding yPC-104 was amplified using error-prone PCR. Oligonucleotides YPC104B and YPC3'C and template pFdBd-104 were employed using the PCR conditions described in 6.2.1, except that dATP was 1 mM and 1 unit of *Taq* DNA polymerase from a batch known to be of low fidelity (ProMega, Batch No. 573 1805) were included in the reaction mix. PCR products were digested with *Nco*I and *Pst*I, and 3.8 pmol of purified insert was incubated with 0.64 pmol of similarly treated pFdBd-K1135L and 2U T4 DNA Ligase (Boehringer Mannheim). DNA was ethanol precipitated in the presence of NaCl and 50 µg carrier tRNA and resuspended in 20µl of water. Electrocompetent XL1-Blue cells were prepared (Sharma & Schimke, 1996) and transformed with 1µl aliquots of the DNA. The transformants were allowed to recover for 1 hour in SOC media (Downer *et al.*, 1988) before an aliquot was withdrawn to estimate the size of the library. Phagemids from the library were prepared using the method of (Gu *et al.*, 1995) with the following alterations: 2YT was supplemented with 100 µg/ml ampicillin and 10 µM biotin, and expression of the gIIIp fusion protein induced with 0.02 mM IPTG. After 1 hour of helper phage infection, 50 µg/ml of kanamycin was added and the culture moved to 30°C overnight. Phagemid particles were isolated by PEG precipitation and stored at -80°C in PBS.

7.2.2 *In vitro* biotinylation of mutant biotin domains on phagemids

Phagemids produced from the library were biotinylated with purified *E. coli* biotin ligase (a kind gift from Dr. D. Beckett, Department of Chemistry and Biochemistry, University of Maryland) in buffer containing 40 mM Tris pH 8.0, 3 mM ATP, 5.5 mM $MgCl_2$, 50 mM KCl, 5 μM biotin, 0.1 mM DTT and 0.1% BSA. Phagemid particles (0.8 - 1.2 x 10^{12} cfu) were reacted with 0.25 pM enzyme at 37°C for varying times before panning. Phagemids were recovered from the reaction by precipitation with PEG and the pellet washed to remove free biotin, before being resuspended in PBS for panning.

7.2.3 Expression and purification of apo yPC-104 and mutant peptides

Overexpression of yPC-104 or mutant peptides was performed in *E. coli* BL21(λDE3) cells using plasmids derived from the pET16b expression vector. These plasmids were constructed by digesting pET16b with *Nco*I and *Pst*I and introducing a similarly treated DNA fragment containing the mutant biotin domain coding regions, excised from the vectors isolated from the *in vivo* selection. Crude cell lysates containing both the apo and holo forms of yPC-104 were prepared using the method of Chapman-Smith *et al.* (1994) except cells were lysed in Buffer A (50 mM Tris-HCl, pH 8.0, 0.1 mM EDTA). The lysate from 1l of culture was filtered and passed through a Q-Sepharose column (Pharmacia, 12 cm x 2.6 cm) equilibrated in Buffer A, and the unbound material containing apo yPC-104 collected. This flow-through fraction was reapplied to the column and fractionated with a 250 ml gradient from 0 to 200 mM NaCl in Buffer A, run at 5 ml/min. Fractions containing apo yPC-104 were determined using SDS-PAGE

and pooled, concentrated and run on a Superdex-75 HR 10/30 (Pharmacia) gel-filtration column in 2 mM ammonium acetate, pH 7.5. Those fractions containing the purified apo-biotin domain were lyophilised and stored at -20°C.

7.2.4 Analysis of biotin domain mutants

Comparative *in vivo* biotinylation of the yPC-104 mutants was performed essentially as described previously by Val *et al* (1995). The mutants were expressed in *E. coli* TM21 cells from the pC-104 derived expression plasmids described above by IPTG induction for 2 hours. Whole cell lysates were prepared using the method described by Chapman-Smith *et al.* (1994). The expressed yPC-104 proteins was resolved from total cellular protein by fractionation on duplicate 15% (w/v) polyacrylamide gels under denaturing and reducing conditions. Protein was visualised by staining with Coomassie Brilliant Blue R250 and biotinylated protein detected by avidin alkaline phosphatase blot (Chapman-Smith *et al.*, 1994). The intensity of both the biotinylated band and the yPC-104 protein band were quantitated by laser densitometry and the extent of *in vivo* biotinylation, expressed as biotinylated protein divided by the total yPC-104 protein, in arbitrary units, for each peptide. This ensured that our quantitation of the extent of *in vivo* biotinylation was unaffected by any variation in expression levels or protein stability resulting from the introduced mutations.

Kinetic analysis of the enzymatic biotinylation was performed, as described in 3.2.3, using purified BPL and apo yPC-104 or mutant peptides.

7.2.5 Molecular Modelling

7.2.5.1 Identification of homologous structures

Protein structures in the PDB predicted to have homology to yPC-104 were identified by the threading method of Fischer & Eisenberg (1996) as implemented on the UCLA structure-prediction server (URL http://fold.doe-mbi.ucla.edu). The protein structures with compatibility (Z) scores higher than the confidence threshold of 5.0 ± 1.0 were: the lipoyl domain of pyruvate dehydrogenase from *Azotobacter vinelandii* [(Berg *et al.*, 1997) IYV, Z=10.89], the biotinyl domain of acetyl-CoA carboxylase from *E. coli* [(Athappilly & Hendrickson, 1995; Roberts *et al.*, 1999) 1BDO, Z = 9.65], the lipoyl domain of *Azotobacter vinelandii* 2-oxoglutarate dehydrogenase [(Berg *et al.*, 1996); GHJ, Z=8.25) and the lipoyl domain of pyruvate dehydrogenase from *Bacillus steorothermophilus* [(Dardel *et al.*, 1993); LAB, Z=5.79]. From the individual sequence alignments, a consensus sequence alignment of yPC-104 to these proteins was generated (Figure 7.3). Due to the lack of N-terminal regions homologous to that of yPC-104 (residues 1075-1098) in the identified structures, only residues 1099-1170 could be modelled. Secondary structure prediction for this region using the PHDsec program (Rost & Saunder, 1993) suggested the presence of an α-helix from residue Lys1077-Val1087 of yPC-104.

7.2.5.2 Homology Model Construction

Homology models of apo yPC-104 and holo yPC-104 residues 1099-1170 were constructed using the HOMOLOGY module of the INSIGHT II software (version 98.0, MSI, San Diego). The structures listed above and the coordinates of the

holo BCCP domain [(Athappilly & Hendrickson, 1995; Roberts *et al.*, 1999); 2BDO] and apo BCCP domain [(Roberts *et al.*, 1999; Yao *et al.*, 1997); 3BDO] were analysed to identify the structurally conserved regions (SCRs) by iterative fitting of C^{α} atom positions and analysis of hydrogen bonding patterns. For NMR ensembles a mean coordinate set, generated using MOLMOL (Koradi *et al.*, 1996) was used in the comparison process. The defined SCRs and loops are indicated in Figure 7.3. For the models of holo and apo yPC-104 the SCRs and loop 1 coordinates were modelled on those of holo and apo BCCP, respectively. Coordinates for loop 2 in each model were based on loop 2 of the lipoyl domain of pyruvate dehydrogenase (1IYV).

7.2.5.3 Refinement of Models

The two models were further refined by restrained simulated annealing using the program X-PLOR (Brunger, 1992) employing the CHARMm force field (Brooks *et al.*, 1983) The restraints employed consisted of the backbone hydrogen bonds from the respective structure upon which the model was based (HN-O 1.7-2.0 Å and N-O 2.7-3.0 Å) and harmonic point restraints on the C^{α} atoms tethering them to their positions in the initial models. This allowed the backbone to remain relatively fixed and the secondary structure to remain intact while side chain positions in the core of the domain could be optimised to reduce steric clashes and promote intimate packing. Simulations were carried out *in vacuuo* with the dielectric constant set to 10 to approximate solvent screening. Coordinate sets were cooled from 2000-100K over 5 ps of dynamics followed by 500 steps of conjugate gradients energy minimisation. For both models ten

simulations were carried out and the five final models with the lowest overall energy were retained for analysis.

7.3 RESULTS

7.3.1 Construction of yPC-104 mutant library

Having established the conditions for phage display and an *in vivo* selection system against yPC-104, a novel method for investigating the protein: protein interaction between a biotin domain and a BPL was devised. Firstly, a library of mutant biotin domains were displayed on the surface of bacteriophage and peptides, which were products of enzymatic biotinylation, were specifically captured with streptavidin. This step allowed the removal from the library those mutations inducing gross structural changes. Following phage display, the *in vivo* selection system was employed to segregate the library into a pool of mutants with decreased affinity for BirA, allowing identification of amino acid residues important for the interaction with the enzyme

A phage display library encoding mutated yPC-104 was created by error prone PCR, using *Taq* DNA polymerase and limiting the availability of dATP in the reaction, as described in 7.2.1. A mini-library, created to assess the efficiency of mutagenesis, was screened by introducing individual PCR products into a cloning vector and sequencing plasmids isolated from 12 colonies. This revealed an average of one point mutation in the 312 base pairs of amplified DNA. A bias of AT to GC mutations was observed but this was not exclusive. After scaling up the library construction and infecting the transformed bacteria with helper phage, a phagemid library of $\approx 7 \times 10^6$ members was generated.

In order to permit mutants of yPC-104 with decreased affinity for biotin ligase to be captured by the panning procedure, library phagemids were incubated with purified BirA *in vitro* before panning. Furthermore, panning was

restricted to two rounds to minimise the possibility of losing low affinity mutants over multiple rounds. For the first round of panning phagemids were reacted with purified enzyme for 24 hours. Panning of the library at this stage yielded a 150 fold enrichment over K1135L control phagemids.

In the second round of panning the time of exposure to the enzyme *in vitro* was varied, with the intention of segregating the library on the basis of the affinity of mutant biotin domains for biotin ligase. Phagemids were treated with purified BirA for 6 hours before being captured by panning. The phagemids in the washes were concentrated and reacted with BirA for a further 24 hours. After panning the phagemids were used to infect XL1-Blue cells and double stranded plasmids were extracted from cells for further analysis using the *in vivo* selection system.

7.3.2 Selection of biotinylation defective mutants *in vivo*

DNA encoding the mutant biotin domains was excised from the phage display vector and cloned into the expression vector pC-104 for the genetic selection. The ligation products were introduced into the biotin auxotroph *E. coli* strain TM21 (Chapman-Smith *et al.*, 1999) and grown on media containing limited biotin, ampicillin and IPTG for selection of biotin domain mutants *in vivo*. Transformants were replica-plated onto both selective and non-selective media for growth phenotype analysis. Approximately 80,000 colonies were screened in this manner and a wide range of colony sizes was observed on the selective media. Since growth under these conditions results from expression of a poorly biotinylated protein (Chapman-Smith *et al.*, 1999), plasmid DNA was isolated from colonies showing the best growth. 18 isolates giving the expected restriction

pattern were further analysed by DNA sequencing, and point mutations were detected in all clones. This data, shown in Figure 7.1, revealed 12 unique single missense mutations. In addition, three silent mutations and two double missense mutations were isolated. Interestingly, a Ser1162→Pro mutation, appeared in both of the segregated pools suggesting that *in vitro* biotinylation had gone to completion within the initial 6 hour reaction. In addition three silent mutations and two double missense mutations were isolated. Constructs harbouring single point mutations from the double mutant His1102→Arg; Met1107→Thr were made to assess the effect of the individual amino acid substitutions. All mutants were further analysed using an *in vivo* biotinylation assay.

7.3.3 Analysis of biotinylation *in vivo* by BirA

The mutants isolated using phage display and the *in vivo* selection (Figure 7.1) were analysed to determine the ability of the peptides to function as substrates for BirA *in vivo* at 37°C. The extent of biotinylation for the mutants, expressed relative to wildtype yPC-104, was quantitated as described in 7.2.4. As expected from previous experiments, the Lys1135→Leu substitution abolished biotinylation. This construct served as a negative control for all experiments. Val residues at three different positions in the biotin domain were substituted with Ala. The substitution at position 1116 did not affect biotinylation in this assay. However, the mutations at positions 1148 and 1166 were more severe causing decreases of 49% and 46% respectively.

A number of other mutations, isolated using this system, were located in the region proposed to form the structured biotin domain. Mutations in this region

of the molecule caused a decrease in biotinylation, with the exception of the His1102→Arg and His1117→Arg substitutions which did not significantly affect biotinylation (Figure 7.1). Two mutations were found at the highly conserved Met1134 residue adjacent to the biotinyl-Lys. Substitution of this residue with Val or Thr caused similar decreases in biotinylation (31% and 26% respectively). Ser residues at position 1141 or 1162 substituted with Pro gave a decrease in biotinylation of 37% and 60% respectively. Phe at position 1152 was substituted with an Ile and caused a decrease in biotinylation to 65%. Two bands were observed on the avidin blot for this mutant suggesting the protein was susceptible to proteolysis by *E. coli* enzymes *in vivo*.

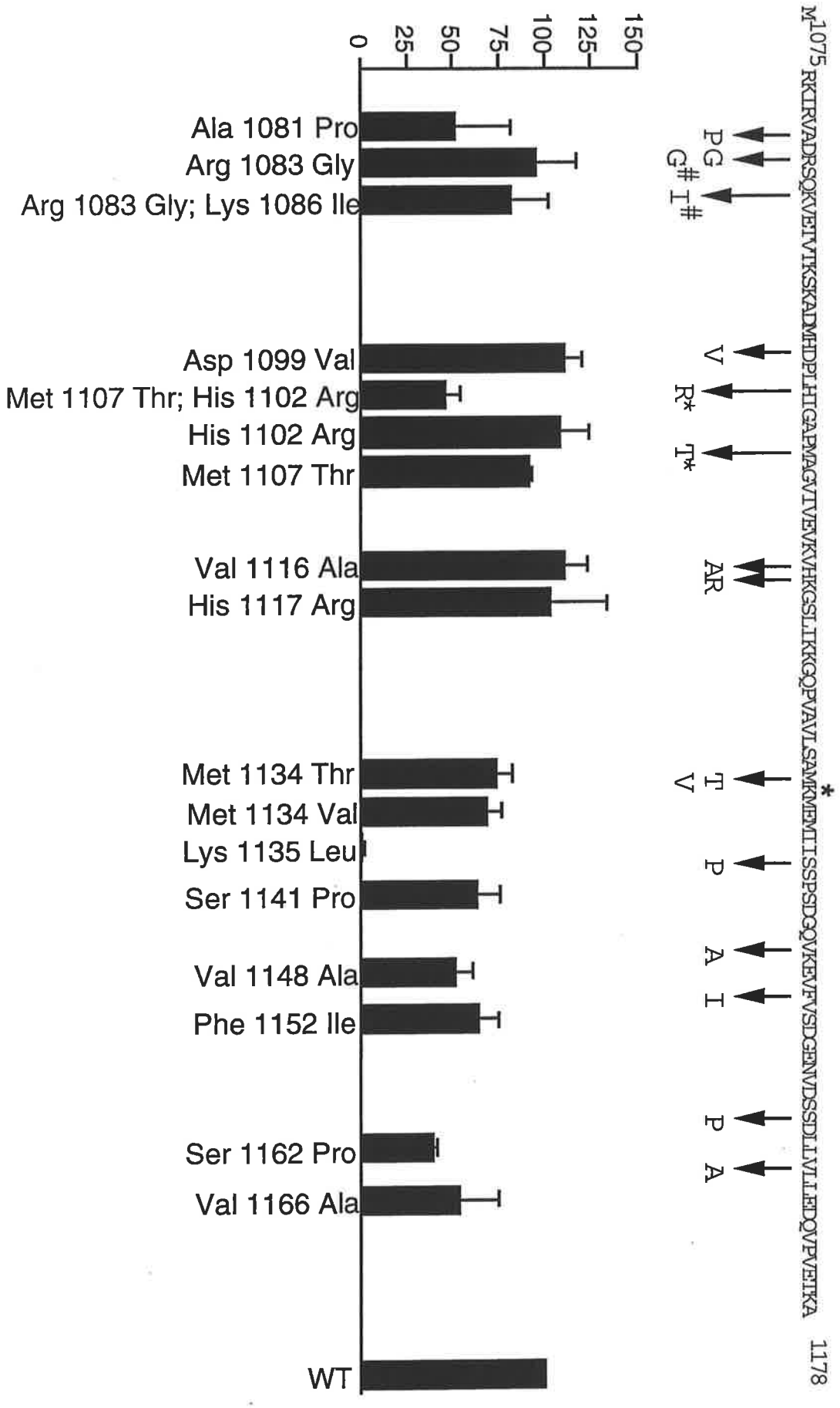
Several mutations were detected in the amino terminal region of yPC-104 which mapped outside the structurally conserved biotin domain. Substitution of Ala1081 with Pro caused a decrease in biotinylation to 52% whereas Arg1083→Gly had only a slight effect on biotinylation (95%). A decrease in biotinylation was observed for the double mutants Arg1083→Gly; Lys1086→Ile (82%) and His1102→Arg; Met1107→Thr (46%). Individual mutations from the His1102→Arg; Met1107→Thr double mutant were cloned into separate vectors and analysed but neither single mutation had a large effect on biotinylation (108% for His1102→Arg, 92% for Met1107→Thr) suggesting that the large decrease observed for the double mutant was synergistic.

7.3.4 Temperature sensitivity of yPC-104 mutants

Temperature sensitive mutations of BCCP have been shown to affect the conformation of the biotin domain and, as a result, substrate recognition

Figure 7.1: *In vivo* biotinylation assay

Biotinylation of yPC-104 mutants at 37°C was analysed *in vivo* in *E. coli* (as described in 7.2.4). The sequence of the 104 C-terminal amino acid residues of yeast pyruvate carboxylase 1 is shown at the top with the biotin accepting lysine residue indicated (*). The position and amino acid substitution of the mutations isolated in this study are indicated below the sequence. Two isolates containing double mutants are represented by # and *. The graph shows biotinylation of the mutants relative to the wildtype peptide.


Relative Biotinylation
(Percentage of Control)
0
25
50
75
100
125
150
Ala 1081 Pro
Arg 1083 Gly
Arg 1083 Gly; Lys 1086 Ile
Asp 1099 Val
Met 1107 Thr; His 1102 Arg
His 1102 Arg
Met 1107 Thr
Val 1116 Ala
His 1117 Arg
Met 1134 Thr
Met 1134 Val
Lys 1135 Leu
Ser 1141 Pro
Val 1148 Ala
Phe 1152 Ile
Ser 1162 Pro
Val 1166 Ala
WT
M1075 RKIRVADRSQKVETVTKSKADMHDPLHIGAPMAGVTVEVKVHKGSLIKKGQPVAVLSAMKMEMTISSPSDGQVKEVFVSDGENVDSSDLLVLLEDQVPVEIKA 1178
PG
G# T#
V
R* T*
AR
T V
P
A I
P A
*

(Chapman-Smith *et al.*, 1999; Li & Cronan, 1992). To investigate the possibility that the amino acid substitutions in yPC-104 induced structural alterations, the *in vivo* biotinylation assay was repeated at a lower temperature. The relative biotinylation of selected mutants was analysed at 30°C and compared with the results at 37°C (Figure 7.2). Substitutions at all positions investigated, with the notable exception of Met1134, showed an increase in the extent of biotinylation at 30°C. Substrates Phe1152→Ile and Val1166→Ala were comparable to the wildtype peptide at the lower temperature (115% and 97% respectively). The proteolytic product of Phe1152→Ile observed at 37°C was not evident at 30°C. Biotinylation increased from 63% to 89% for Ser1141→Pro, from 51% to 67% for Val1148→Ala and from 40% to 67% for Ser1162→Pro. These data suggest that point mutations of yPC-104 which were temperature sensitive most likely altered the conformational stability of the biotin domain rendering it a less favourable substrate for BirA at the higher temperature. In contrast, substitution of Met1134 with either Val or Thr were not temperature sensitive mutations. Biotinylation of Met1134→Thr was the same at either temperature (74%) whereas Met1134→Val decreased only slightly from 69% at 37°C to 62% at 30°C. These two temperature insensitive mutations were chosen for further analysis.

7.3.5 Predicted structure of yPC-104

To assist in interpreting the results obtained from *in vivo* analysis of apo yPC-104 mutants, we constructed a molecular model of the yeast biotin domain. This model was based on the solution structures of apo BCCP-87 and several known structurally analogous lipoyl domains, as described in 7.2.5. The

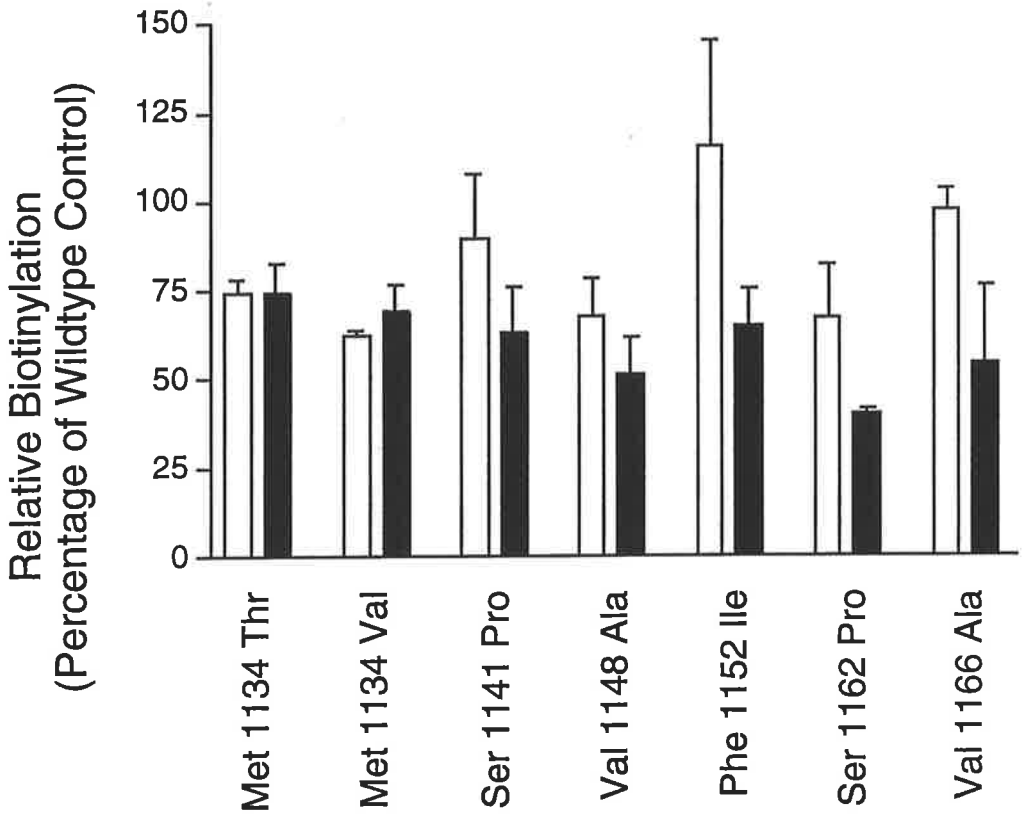


Figure 7.2: Temperature sensitivity of yPC-104 mutants

The *in vivo* biotinylation assay described in Figure 7.1 was repeated at 30°C (white) and 37°C (black). The graph represents biotinylation of the mutants relative to the wildtype peptide at each temperature.

predicted structure of holo yPC-104 was also performed, modelled upon holo BCCP, with the two structures being essentially identical. For comparison purposes, the first member of the 3BDO (apo BCCP) and 2BDO (holo BCCP) NMR ensembles were subjected to identical "refinement" protocols. For yPC-104 and BCCP, ten coordinate sets were generated for each model and the five best kept. Measures of coordinate deviations, energies and stereochemical properties of the yPC-104 and BCCP model ensembles are shown in Table 7.1. All the model ensembles are tight, in terms of atomic rmsd values and energetics, with any one model being as good as any other. The deviation of model BCCP coordinates from the starting experimental coordinates is similar to the rmsds for the NMR ensembles (for all backbone atoms: apo BCCP, models: 1.78 ± 0.03 Å, NMR: 1.57 ± 0.48 Å; holo BCCP, models: 1.04 ± 0.02 Å, NMR: 1.33 ± 0.24 Å). The BCCP models have markedly improved Ramachandran properties (apo BCCP; models: 69.6 %, NMR: 40.0 %; holo BCCP, models: 74.9 %, NMR: 50.5 % of residues in the most favoured regions). The yPC-104 models also exhibit comparable backbone atom deviations from their BCCP-based starting coordinates (apo yPC-104: 1.26 ± 0.04 Å and holo yPC-104: 1.17 ± 0.04 Å) and have energy values comparable to those of the BCCP models for all terms except the harmonic term and the electrostatic term. The difference in the magnitudes of harmonic terms between the yPC-104 and BCCP models is interesting and suggests that the BCCP backbone atom positions suit the yPC-104 models very well (better than the BCCP models) but as the harmonic term has no physical basis and is imposed to stop the models "wandering" too far from the experimentally determined positions this difference is of little significance. The electrostatics of the BCCP models are better than those of the yPC-104 models, which is not surprising as the charged side chain positions in the BCCP models

Table 7.1: Measures of coordinate deviations, energies and stereochemical properties of the models and the NMR ensembles

Energies[1]	apo yPC	holo yPC	apo BCCP	holo BCCP
Overall	53 ±14	80 ± 9	170 ± 6	128 ± 7
Bonds	22.4 ± 0.3	22.6 ± 0.7	30.2 ± 0.3	30.7 ± 1.1
Angles	122 ± 3	125 ± 1	147 ± 5	141 ± 4
Dihedrals	258 ± 3	264 ± 4	319 ± 8	311 ± 6
Impropers	3.7 ± 0.1	3.4 ± 0.1	3.7 ± 0.3	3.2 ± 0.2
VDW	-349 ± 7	-344 ± 4	-348 ± 4	-351 ± 4
Distance	0.23 ± 0.02	0.30 ± 0.07	0.09 ± 0.02	0.17 ± 0.08
Elec	-8.6 ± 3.2	5.9 ± 1.3	-26.5 ± 2.3	-41.6 ± 6.6
Harmonic	4.4 ± 1.1	3.8 ± 0.9	44.9 ± 2.6	33.5 ± 3.0
B.B. rmsd[2]	0.27 ± 0.10	0.27 ± 0.11	0.30 ± 0.07	0.34 ± 0.07
H.A. rmsd[3]	0.63 ± 0.17	0.66 ± 0.20	0.75 ± 0.17	0.86 ± 0.20

Ramachandran Properties[4]				
Region	apo yPC	holo yPC	apo BCCP	holo BCCP
Most Fav.	67.7	72.0	69.6	74.9
Allowed	26.0	23.7	23.8	19.1
Generous	3.7	3.7	4.3	3.6
Disallowed	2.6	0.6	3.2	2.4

[1] Energy terms in kcal.mol^{-1} calculated in X-PLOR using default CHARMm22 parameters and with a dielectric constant of 10. The Distance term refers to the violation of hydrogen bond restraints (treated as NOE restraints) not hydrogen bond energies which are implicit in the VDW and Elec terms.

[2] Backbone (N,C$^{\alpha}$,C') pairwise rmsd of the 5 refined models.

[3] Heavy (non-hydrogen) atom pairwise rmsd of the 5 refined models.

[4] Calculated for each residue in each conformer (not angular averages) using the in-house program ANGORDER. Values are percentages of non-Gly and non-Pro residues and exclude the N- and C-terminal residues.

have some experimental basis. In addition, yPC-104 has much higher charge density on its surface and thus the potential for many more unfavourable as well as favourable interactions.

From sequence alignments, shown in Figure 7.3, the eight β-strands of apo yPC-104 were defined as structurally conserved regions (β1 Leu 1101-Ala1105, β2 Val1110-Ile1111, β3 Leu1121-Ile1122, β4 Pro1127-S1132, β5 Glu1137-Ser1141, β6 Gly1146-Val1151, β7 Glu1157-Val1159, β8 Asp1163-Leu1168). Like other biotin and lipoyl domains, the molecule is stabilised by a central hydrophobic core. In addition, a hydrophobic patch exists on the surface of the molecule formed by the side chains of Val1151, Phe1152, Val1153 and Leu1164.

7.3.6 Interpretation of mutational analysis

In this study, temperature sensitivity was used to assess the structural stability of yPC-104 mutants. Of the three Val→Ala mutants isolated, two were investigated for temperature sensitivity. These three Val residues, Val1116, Val1148 and Val1166, contribute to the hydrophobic core of the predicted structure (Figure 7.4) and are involved in forming contacts between each other. It is reasonable to conclude that although mutation to Ala is conservative, replacement of the larger Val side chain with Ala distorts the packing arrangement of the hydrophobic core decreasing protein stability at the higher temperature. The hydrophobic side chain of Phe1152 is not buried in the core but is exposed on the surface of the molecule (Figure 7.5) . In the crystal structure of BCCP the corresponding residue (Leu139) interacts with another hydrophobic

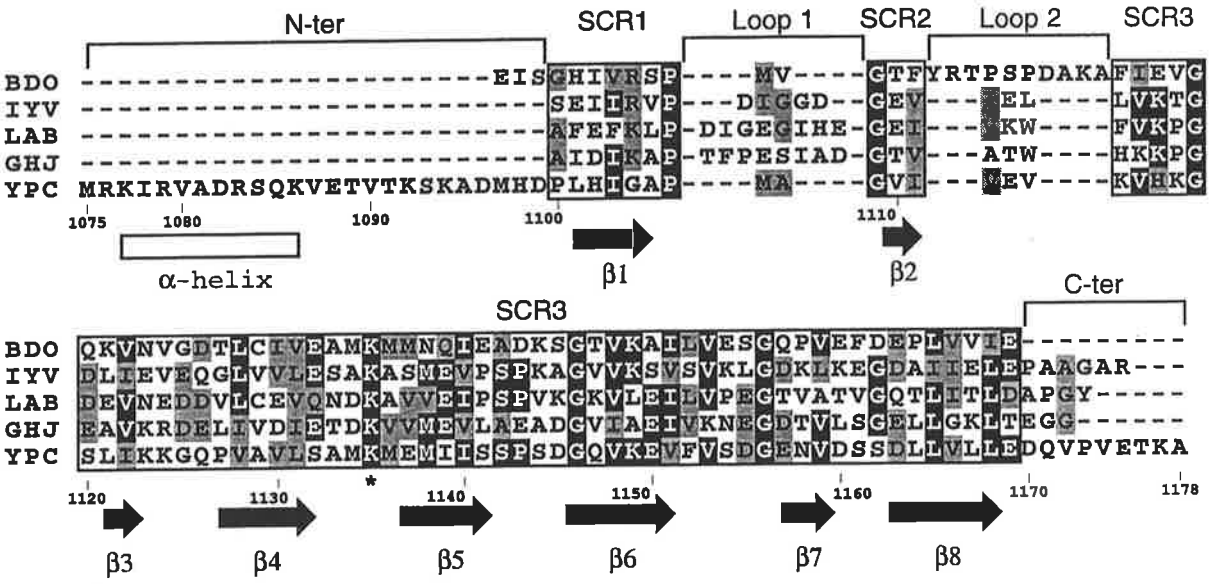


Figure 7.3: Sequence alignments of structurally related biotin and lipoyl domains

The primary sequences of the biotin domain of *E. coli* acetyl CoA carboxylase (BDO) and the lipoyl domains of *A. vinelandii* pyruvate dehydrogenase (IYV), 2-oxoglutarate dehydrogenase (GHJ) and *B. stearothermophilus* pyruvate dehydrogenase (LAB) were aligned with the 104 C-terminal residues of yeast PC (yPC), as described in 7.2.5.1. Amino acid residues conserved between the domains are coloured white on black background. Residues similar to those found in the *E. coli* sequence at analogous positions are shaded. The structurally conserved regions (SCR) and non-conserved loop regions are indicated above the alignment. The amino acids proposed to form secondary structure in the yPC model are shown under the alignment, as are the positions of residues present in yPC. The biotinyl- and lipoyl- lysines are indicated (*).

residue (Ile78) found in the short N-terminal extension outside of the structured domain. It is likely that under the appropriate conditions this hydrophobic region is involved in a packing interaction with other parts of the molecule and the change to Ile disrupts this interaction.

The modelling suggests that the two temperature sensitive Ser→Pro mutations may introduce local conformational changes which render the domain a less favourable substrate for BirA. The peptide backbone angles of Ser1162 (ϕ = +65°) are unfavourable for introducing a Pro residue at this position (fixed $\phi \approx$ -75°). Although a Pro substitution at Ser1141 (ϕ = -94°) maybe tolerated in this regard, other factors may compromise the structure with this mutant. It is possible that the Pro substitution may abolish a potential interaction between the side chain of Ser1141 and the backbone carbonyl oxygen of Gly1125 which are approximately 2 Å apart in the model. Furthermore, the dihedral angles, and subsequently hydrogen bonding interactions between the β4 and β5 strands, may also be affected by the substitution. These strands define the hairpin loop containing the biotin-accepting lysine residue. Met1134 resides on this hairpin loop, together with Lys1135, and does not form backbone hydrogen bonds with other residues. Therefore substitution of Met1134 with either Val or Thr is unlikely to destabilise this hairpin turn. The observed temperature sensitivity of the mutants that were isolated is consistent with their proposed contribution within the structured biotinyl domain of yPC-104.

The secondary structure of the 23 amino acids N-terminal amino acids of yPC-104, lying outside of the structured biotin domain, was predicted using the PHDsec programme (Rost & Saunder, 1993). Residues Lys1077 to Val1087,

were predicted to form an α-helix (Figure 7.3). The point mutations Ala1081→Pro and Arg1083→Gly obtained in the mutational analysis of yPC-104 mapped to residues forming this α-helix. Both substitutions introduced residues known to disrupt α-helices, suggesting that this region of the molecule may indeed be structured. It has been previously shown that expression of a peptide lacking the 23 N-terminal amino acid residues is a 6-fold less favourable substrate in *in vivo* biotinylation assays than the full length, 104 amino acid residue peptide (Val *et al.*, 1995) In addition, the 2D TOCSY NMR spectra, performed using unlabelled protein, revealed that yPC-104 is structured whereas the shorter peptide is not (E. L. Roberts, A. Chapman-Smith, R. N. Perham and J. C. Wallace, unpublished data). Together these data suggest that the N-terminal region is structured and contributes to the stability of the molecule.

The surface mutation of Phe1152→Ile showed temperature sensitivity in the *in vivo* biotinylation assays. This suggests that the hydrophobic patch on the surface of yPC-104, containing Phe1152, plays a role in the stability of the domain (Figure 7.4 and 7.5). Furthermore, it was observed that this mutant was susceptible to proteolysis during the assays, supporting the involvement of Phe1152 in an interaction stabilising the domain. It is possible that the hydrophobic patch on the surface of yPC-104 may interact with the N-terminal region of yPC-104, possibly through a hydrophobic patch on the predicted α-helix.

Figure 7.4: Predicted structure of the yeast PC biotin domain

A schematic diagram of the predicted structure of the yeast PC biotin domain, produced from the molecular modelling described in section 7.2.5. The position of the biotin accepting lysine at position 1135 and the adjacent Met1134 are shown, as are the side chains of residues which, when mutated, introduced temperature sensitivity into the structure. The figure was generated using Molscript (version 2.1.1, Kraulis, 1991) and Raster3D (version 2.4, Merritt & Bacon, 1997).

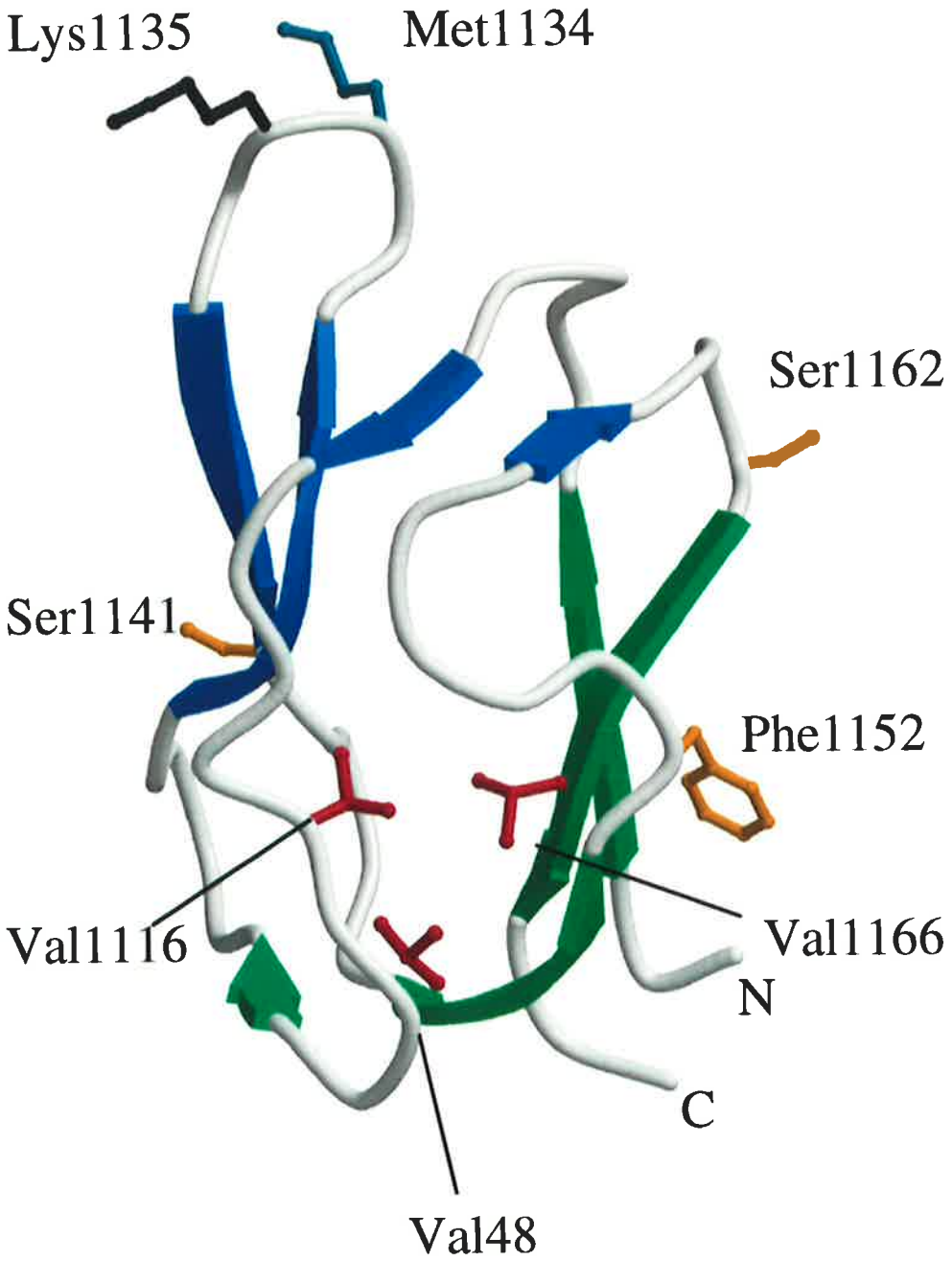
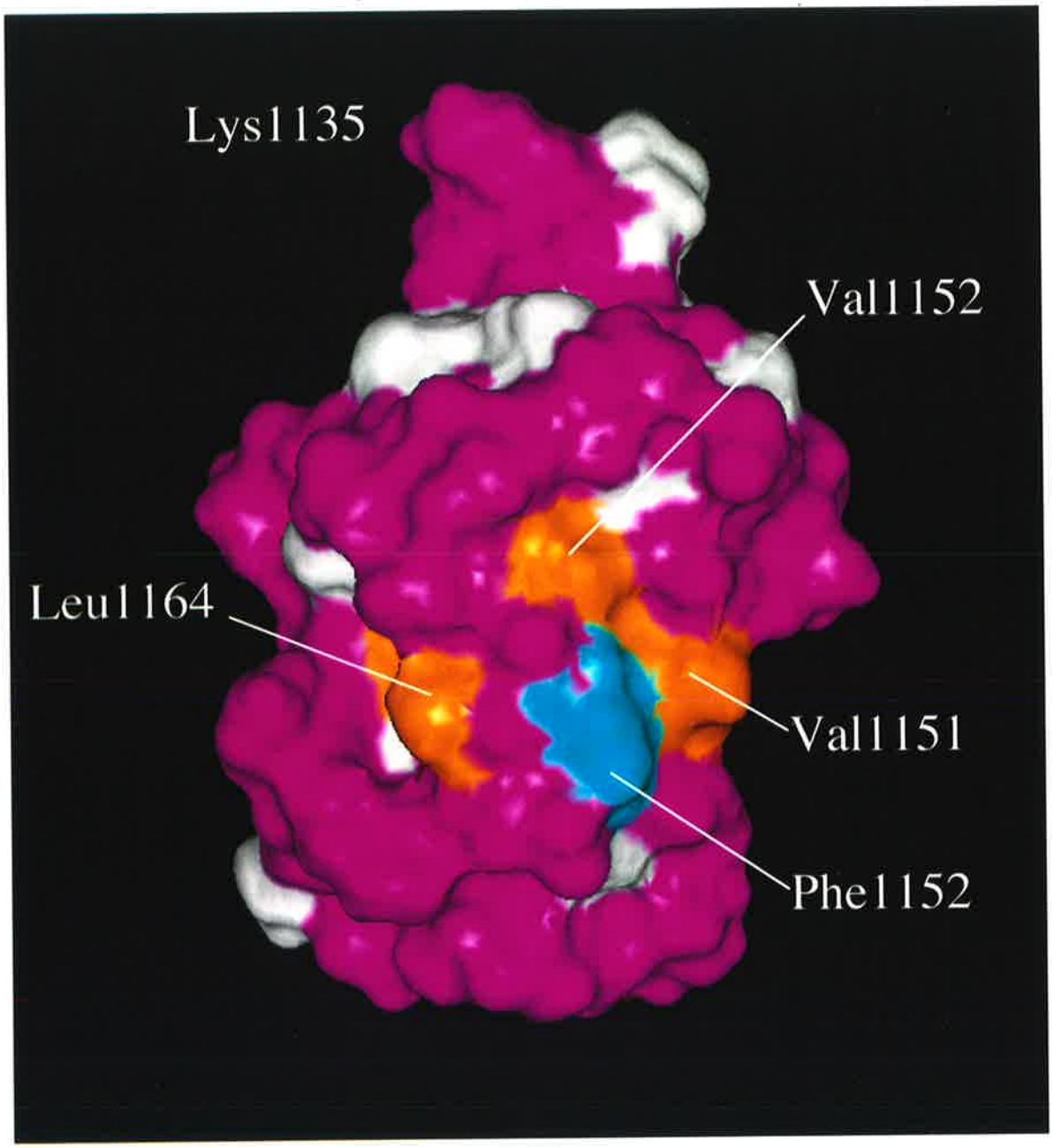

Lys1135
Met1134
Ser1162
Ser1141
Phe1152
Val1116
Val1166
N
C
Val48

Figure 7.5: Hydrophobic patch on the surface of the yPC-104 predicted structure.

A space-filling representation of the yeast PC biotin domain predicted structure. The hydrophilic surfaces are coloured magenta while buried hydrophobic residues are coloured white. The hydrophobic side chains of Phe1152 (coloured blue) and Val1151, Val1153 and Leu1164 (orange), present on the surface of the domain, are shown. The position of the biotin accepting lysine at position 1135 is at the top of the figure. The figure was generated using Raster3D (version 2.4, Merritt & Bacon, 1997).


Lys1135
Val1152
Leu1164
Val1151
Phe1152

7.3.7 Kinetic Analysis of Biotinylation *in vitro*

The peptides yPC-104, Met1134→Val and Met1134→Thr were purified as unbiotinylated domains. Analysis of the peptides by electrospray mass spectroscopy revealed that the purified peptides were 131 mass units less than the expected molecular mass for the non-biotinylated form, suggesting the N-terminal Met was excised *in vivo.* (Table 7.2).

Table 7.2 : Mass spectrometric analysis of the purified yPC biotin domains

Peptide	Calculated apo protein mass	Calculated apo protein mass minus Met[1]	Mass determined
yPC-104	11452.4	11321.2	11322
Met1134→Val	11420.3	11289.1	11289
Met1134→Thr	11422.3	11291.1	11292

The ability of these biotin domains to function as substrates for BirA was assessed using steady state kinetics by varying the concentration of the apo biotin domains whilst keeping the concentration of biotin and MgATP at constant, saturating levels. The assays were performed under conditions that were optimal for biotinylation of BCCP-87 (Chapman-Smith *et al.*, 1999) Kinetic constants derived from these data are shown in Figure 7.6. BCCP-87 was the best substrate for the bacterial enzyme displaying a K_m of 3.2 μM. The k_{cat}/K_m values of the different substrates revealed that the affinity of the wildtype yeast biotin

domain for BirA was approximately 10-fold lower than BCCP-87. Similar k_{cat} values were observed for these two substrates, suggested that the deficiency in biotinylation is not due to changes in the rate of catalysis. As expected, the two yeast domain mutants were poor substrates for BirA. Biotinylation could only be detected when the concentration of the enzyme in the assay was increased 10-fold, and sufficiently high concentrations of the mutants were used. Both mutants displayed a K_m at least 100-fold greater than BCCP-87 and 10-fold higher than yPC-104.

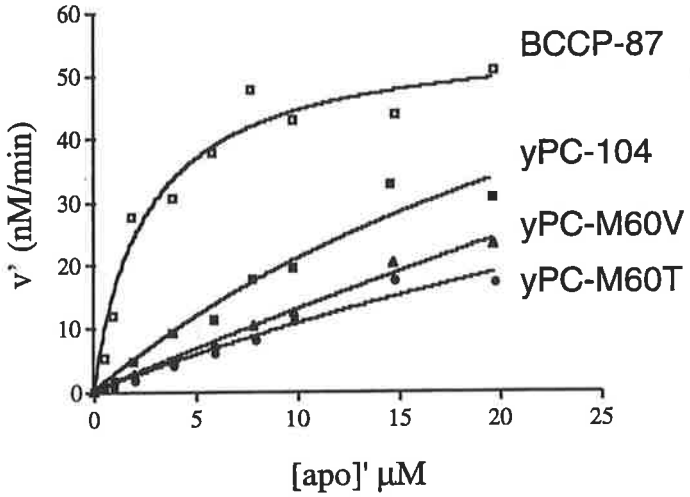
The mutants were also assayed with yeast biotin ligase under conditions which were optimal for biotinylation of yPC-104 (Chapter 3). The wildtype yeast biotin domain was the preferred substrate for yBPL (Figure 7.7), with the affinity for BCCP-87 being 10-fold lower. The Met1134→Val mutant, which was a poor substrate for BirA, displayed only a 2-fold lower affinity than the wildtype yPC-104. Again, similar rates of biotin transfer onto these peptides (ie k_{cat}) were observed. Thus the differences in the calculated k_{cat} /K_m values arise from the differences in the efficiency of association between enzyme and protein substrate. As was observed with BirA, Met1134→Thr was also a poor substrate for the yeast enzyme displaying 25-fold lower affinity than wildtype.

7.3.8 Trypsin Susceptibility.

In vivo biotinylation of Met1134→Thr and Met1134→Val at 30°C and 37 °C suggested that these substitutions had little effect on the structure of the domain. As a second probe for structural alteration in these mutants, the purified apo proteins were subjected to limited trypsin digestion. The susceptibility to tryptic

Figure 7.6: Kinetic analysis of yPC mutants with *E. coli* BirA

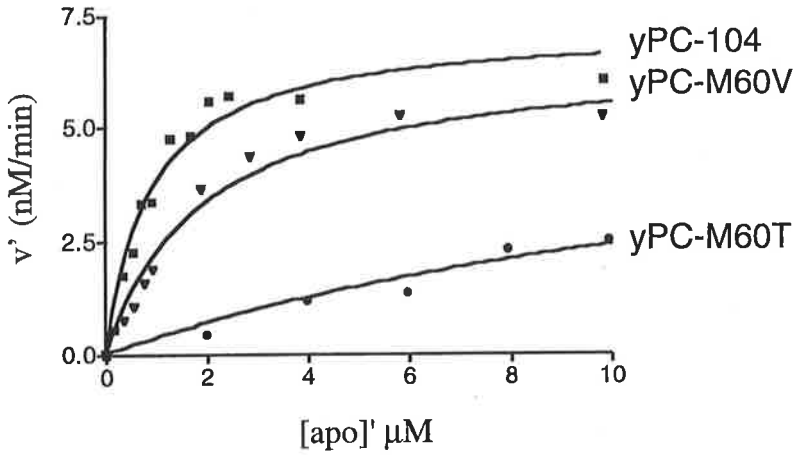
Purified biotin domains were assayed with BirA using the conditions described by Chapman-Smith *et al.* (1999). The activity of the enzyme was determined over the concentration range indicated on the graphs above. In the graph shown, 10-fold more BirA was included in the reactions for the two yPC mutants. The kinetic constants derived from the graphs are shown in the table. The values given represent the mean and S.D. for at least three experments. The K_m values for the two mutant peptides were estimated from a double reciprocal plot. N/D, not determined.



Mutant	k*cat* (s^{-1})	K*m* (μM)	k*cat*/K*m* x10^3 $M^{-1}s^{-1}$
BCCP-87	0.092 ± 0.02	3.2 ± 0.5	28.5 ± 0.01
yPC-104	0.085 ± .011	27.2 ± 5	3.2 ± 0.35
yPC-M60V	N/D	320	N/D
yPC-M60T	N/D	> 320	N/D

Figure 7.7: Kinetic analysis of yPC mutants with yeast BPL

Purified biotin domains were assayed with yeast BPL using the conditions described in section 3.2.3. The activity of the enzyme was determined over the concentration range indicated on the graphs above. The kinetic constants derived from the graphs are shown in the table. The values given represent the mean and S.D. for at least three experments. The K_m value for Met1134→Thr was determined from a double reciprocal plot. N/D, not determined.



Mutant	**k*cat* (s^{-1})**	**K*m* (µM)**	**k*cat*/K*m* x10^5 $M^{-1}s^{-1}$**
yPC-104	0.664 ± 0.1	1.0 ± 0.2	6.63 ± 0.4
yPC-M60V	0.594 ± 0.16	2.0 ± 0.2	2.89 ± 0.6
yPC-M60T	0.372 ± 0.2	13.4 ± 2.8	0.27 ± 0.1
BCCP-87	0.739 ± 0.1	11.1 ± 1.0	0.66 ± 0.1

digestion has previously been used to investigate the effect of point mutations on the structural integrity of a biotin domain (Chapman-Smith *et al.*, 1999). Initially the digestion profile of both the apo and holo peptides was investigated by monitoring the reaction with SDS-PAGE (Figure 7.8). Both substrates produced similar digestion products but the holo domain appeared to be less sensitive to cleavage than the unbiotinylated peptide. The biotin domains were rapidly cleaved within the first 5 minutes of the reaction yielding a product with an apparent molecular weight of 11 kDa. A second, slower cleavage then occurred generating a stable product which migrated at around 10 kDa. For apo yPC-104, this cleavage product was initially detected after 30 minutes of trypsin treatment and constituted approximately 50% of the products at 1 hour and greater than 95% of the products at 2 hours. The cleavage of the 11 kDa product from holo yPC-104 was slower, as after 2 hours approximately half of the sample was still the 11 kDa fragment. N-terminal sequencing revealed two scissile bonds between Arg1083-Ser1084 and Lys1086-Val1087 both in the N-terminal extension, at or near the C-terminal end of the proposed α–helix. Cleavage in this region yielded a product which was relatively resistant to further proteolysis as it was the major species even after 5 hours of treatment of both apo and holo yPC-104. Together these data suggest that the C-terminal region of yPC-104 containing the structured biotin domain was relatively resistant to treatment with trypsin

The two purified apo mutant biotin domains were treated with trypsin using the conditions employed for the wildtype peptide. Analysis of the products by SDS-PAGE revealed that both the mutant peptides were cleaved at the same rate as wildtype protein and generated the same products (Figure 7.8). These

results suggest that substitution of Met1134 with either Thr or Val did not induce significant conformational changes to the structured domain.

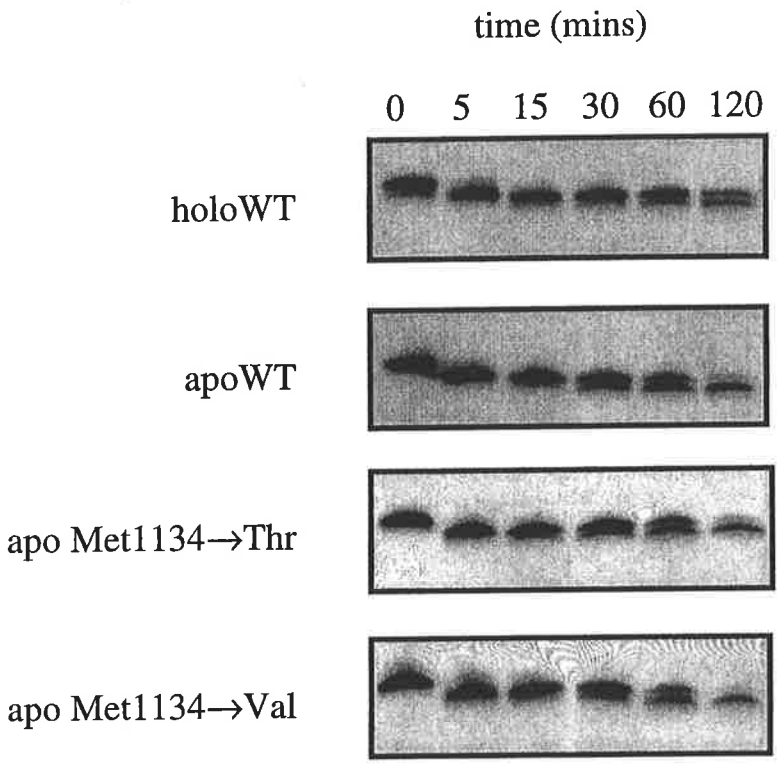


Figure 7.8: Tryptic digestion of yeast biotin domains

Proteins were treated with trypsin (1:40 (w:w) trypsin to protein ratio) in 50 mM Tris-HCl, pH 8.0, and 5 mM $CaCl_2$. Samples, 2.5 µg, were taken at various time intervals indicated above and analysed by SDS-PAGE. The panels show the digestion products of holo yPC-104, apo yPC-104, apo Met1134→Thr and apo Met1134→Val.

7.4 DISCUSSION

In all biotin-dependent enzymes, the residues immediately flanking the biotin-accepting Lys are invariantly Met. Mutation of these residues in both the biotin-accepting subunit of *P. shermanii* transcarboxylase (Shenoy *et al.*, 1992) and in synthetic peptides (Kondo *et al.*, 1984) has revealed that these Met residues are required for the carboxylation of biotin, an event essential for the enzymatic function of the biotin enzymes. It has been demonstrated that the flanking Met residues are not essential for biotinylation by BPLs since conservative substitutions are still recognised as substrates (Shenoy & Wood, 1988; Leon-Del-Rio & Gravel, 1994). In this current work two classes of mutations to the biotin domain of pyruvate carboxylase from *Saccharomyces cerevisiae* were isolated by screening a library of randomly mutated polypeptides using phage display and a genetic selection in *E. coli* that permits the isolation of peptides with a lowered affinity for bacterial BPL. The first class of mutations isolated contained two distinct, conservative substitutions of the Met residue immediately N-terminal of the target Lys. One mutant, Met1134→Val, was a poor substrate for both the bacterial and yeast BPLs. In contrast, although the Met1134→Thr substitution was also a poor substrate for the bacterial BPL, it was only a 2-fold poorer substrate for the yeast BPL than the wildtype protein. Our data show that there are subtle differences in the active sites of BPLs from various species which tolerate mutations to the flanking Met to varying degrees. Furthermore, it has been known for some time that BPLs display cross species reactivity (Chapman-Smith & Cronan, 1999a&b). Here we provide the first kinetic evidence that BPLs biotinylate heterologous biotin domains with a lower efficiency than they do their native substrates. Both the yeast and bacterial BPLs

displayed at least a ten-fold greater affinity for their homologous substrates over heterologous biotin domains (Figures 7.6 and 7.7). Together these data indicate that although BPLs catalyse the same reaction, there are distinct differences between the two enzymes for optimal substrate recognition.

It is apparent that BPLs recognise and interact with a structured biotin domain (Reche *et al.*, 1998; Chapman-Smith *et al.*, 1999). The second class of mutants isolated using our selection technique mapped to residues some distance from the target lysine. All mutants analysed belonging to this class were temperature sensitive, suggesting that these amino acid substitutions induced conformational changes to the folded domain making it a compromised substrate at 37°C. This finding is consistent with our previous work which has demonstrated that several mutations located at a distance from the target lysine of the *E. coli* BCCP biotin domain reduce biotinylation as a consequence of structural alteration (Chapman-Smith *et al.*, 1999). To facilitate interpretation of our mutant data, a molecular model of yPC-104 was produced based upon known biotin domain structures. Our models of yPC-104 differ from that described by Brocklehurst & Perham (1993) primarily due to the availability of additional experimentally determined structures from which to construct a more optimal model of the protein. Using the BCCP coordinates as a guide allowed us to better model the loop between β-strands 1 and 2 and thus our positioning of β-strand 1 differs in register by two residues. We also found no requirement for lengthening the loop between β-strands 7 and 8 and thus our alignment of β-strand 8 differs in register by one residue position.

The mutations obtained in the present study were mapped onto our model. All the temperature sensitive mutations, predicted to induce conformational

changes to the molecule, were at residues implied to be necessary for defining protein structure. In addition, the molecular modelling of the yeast biotin domain highlighted some novel properties, especially in the N-terminal region outside of the structured domain. This region contains a predicted α-helix with three hydrophobic residues (Ile1078, Val1080 and Ala1081) which may function as a cap, packing against the surface of the biotin domain via the hydrophobic surface involving residues Val1151, Phe1152, Val1153 and Leu1164. We have previously shown that expression of a peptide lacking the 23 N-terminal amino acid residues is a 6-fold less favourable substrate in *in vivo* biotinylation assays than the 104 residue peptide (Val *et al.*, 1995). Without the N-terminal extension, the domain was unstable during purification and displayed a marked tendency to form high molecular weight aggregates on gel filtration chromatography (A. Chapman-Smith, and J. C. Wallace, unpublished data). These observations are consistent with our structural predictions and mutant data, which suggest that the N-terminal extension plays a role in stabilising the overall domain structure, rendering the molecule a better substrate for BirA.

The BCCP domain contains a protruding "thumb" between strands β2 and β3 (Athappilly & Hendrickson, 1995; Roberts *et al.*, 1999) that is not present in the structurally heterologous lipoyl domains. Interestingly, this "thumb" contacts the biotin moiety in both the crystal (Athappilly & Hendrickson, 1995) and solution (Roberts *et al.*, 1999) structures of BCCP. The sequence of amino acids which form this thumb is absent in most other biotin domains including *P. shermanii* TC (Chapman-Smith & Cronan, 1999a&b). Furthermore, determination of the structure of the 1.3S subunit of *P. shermanii* TC has revealed that no thumb structure is present in the domain (Reddy *et al.*, 1998) and the biotin moiety does

not contact the protein (Reddy *et al.*, 1997). In this regard, our predicted fold of yPC-104 more closely resembles the structure of the 1.3S subunit of TC. It is possible that the absence or presence of the "thumb" on the biotin domain contributes to the 10-fold reduction in affinity shown by the bacterial and yeast BPL with a heterologous substrate.

Biotinylation and lipoylation *in vivo* by BirA and the two lipoyl attachment enzymes, LplA and LipB, respectively occur in a highly specific manner despite the striking structural similarity of the accepting domains (Wallis & Perham, 1994; Reche & Perham, 1999). It has been demonstrated that the thumb structure of BCCP functions to inhibit lipoylation, as a thumb deletion mutant of BCCP-87, which still functioned as a substrate for BirA, was lipoylated *in vivo* and *in vitro* by LplA (Reche & Perham, 1999). The lipoyl domains lack the thumb, having three residues which span the region corresponding to the base of the thumb, but contain an extended surface loop between β-strands 1 and 2 which is close in space to the lipoyl-lysine (Berg *et al.*, 1996; Berg *et al.*, 1997; Dardel *et al.*, 1993). The residues forming the surface loop are not required for *in vivo* lipoylation, but instead are necessary for the reductive acetylation of the prosthetic group (Wallis *et al.*, 1996). However, as biotinylation of the lipoyl domain does not occur in *E. coli in vivo* (Wallis & Perham, 1994), this surface loop may also function to inhibit recognition by bacterial BPL. In the BCCP domains this middle finger loop is replaced by two residues spanning the region. The yPC-104 domain lacks both the thumb and the extended loop, further suggesting that structural differences govern substrate recognition by yeast and bacterial BPLs.

The selection procedure we used was established using *E. coli* BCCP-87 (Chapman-Smith *et al.*, 1999). We showed that despite BirA having a 10-fold

higher affinity for its native substrate over yPC-104 *in vitro*, high level expression allowed the *in vivo* selection to be adapted for mutational analysis of the yeast protein. Although no wildtype DNA was isolated after selection, several silent mutations were obtained. Mutations which destabilise mRNA, reduce translation efficiency or interfere with protein stability would all decrease the selection pressure on the host. These aspects of our selection procedure account for those mutants obtained in the selection that functioned as wildtype-like substrates in the *in vivo* biotinylation assay, since this assay corrects for variations in expression levels. It should be noted that the relative affinities of the mutant peptides obtained from the *in vivo* and *in vitro* biotinylation assays are a reflection of the different assay conditions. In the *in vivo* assay which measures accumulated product, factors such as the concentration of enzyme, ATP and biotin cannot be controlled. On the other hand, the *in vitro* biotinylation assay measures initial velocity under steady-state conditions using purified components with optimal, saturating concentrations of ATP and biotin.

The enzymatic reaction catalysed by BPLs proceeds through two partial reactions. In the first partial reaction, ATP and biotin bind BPLs in an ordered manner (Xu & Beckett, 1997; Polyak *et al.*, 1999) and the carboxyl group of inert biotin is activated by the addition of an adenylate group. In the second partial reaction nucleophilic attack upon activated biotin by the amine group of the Lys in the protein substrate results in the transfer of biotin from the adenylate onto the apo-biotin domain, with AMP behaving as the leaving group (Xu & Beckett, 1997). The Lys side chain must be precisely positioned into the active site of BPL to bring the electron donating nitrogen close enough to the reactive biotinyl-5'-AMP moiety to permit the chemical reaction. Thus, the biotin domain functions as a protein scaffold that presents the biotin-accepting Lys to the BPLs. The position

of the target Lys in the exposed β-turn between β-strands 4 and 5 is essential in this process, as moving this residue by one place either side abolishes biotinylation (Reche *et al.*, 1998). The kinetic data presented here suggest that it is primarily the protein:protein interaction that is compromised by mutations to this scaffold rather than the catalytic step, since K_m values were affected to a greater extent than k_{cat} values. Although direct measurement of the binding interactions would address this point more fully, it is technically difficult to do so. The form of BPL that interacts with the apo biotin domain is the biotinyl-AMP-BPL complex, and once associated with an unbiotinylated substrate, this complex turns over very rapidly to produce the holo-biotin domain and free enzyme. To undertake such experiments requires either a non-hydrolysable analogue of biotinyl-AMP. or a mutant biotin domain where the target Lys has been modified such that it cannot accept biotin. We have produced a BCCP-87 mutant where the target lysine has been substituted with Leu, but this appears to be a poor inhibitor in studies using BirA (Chapman-Smith, A., Mortellaro, S., Polyak, S. W., Cronan, J. & Wallace, J. C., unpublished data). It is possible that BPL recognition absolutely requires the presence of a target Lys in a structured biotin domain.

Chapter 8

FINAL DISCUSSION

FINAL DISCUSSION

8.1 RECOMBINANT YEAST BIOTIN LIGASE

In Chapter 3, I have discussed the development of a system for the expression and purification of recombinant yBPL. In work presented in this thesis, the purified material was used for domain mapping and in enzyme assays for the kinetic analysis of substrate binding. Collaborations currently being established between the laboratories of Dr John Wallace, Dr Brian Matthews and Dr Dorothy Beckett will investigate the structure of yBPL using X-ray crystallography. This procedure typically requires large quantities of highly pure material. The fusion of a hexahistidine sequence on the C-terminus of the enzyme has facilitated the purification of the recombinant protein using nickel chelating chromatography. Ion exchange was employed to remove a number of contaminants from the initial purification step and gel filtration isolated full length yBPL from material cleaved by a bacterial protease. In order to optimise recovery of the large quantities of yBPL required for structural studies, some refinements to the system developed here should be considered.

Firstly, the limiting step in the production of large quantities of yBPL appears to be the use of *E. coli* as a host for protein expression. It was observed that inclusion body formation accompanied overexpression of yBPL in bacteria. A lower level of expression was accepted in order to improve protein solubility. Yeast is an obvious replacement for bacteria as a recombinant expression system. A number of suitable expression systems have been constructed for

large scale protein expression. A commonly used example is the pYES system (Invitrogen) which permits high level, galactose inducible expression of recombinant proteins. In addition, yeast strains contain the chaperonins necessary to correctly fold yBPL. The enzyme is quite large (76 kDa) relative to the average size of an *E. coli* protein (35 kDa) (Netzer & Hartl, 1997). Eukaryotic cells possess efficient cotranslational folding mechanisms which allow folding of large multidomain proteins *in vivo* (Ellis & Hartl, 1999). Thus for large scale yBPL production, a yeast expression system should be trialed.

Secondly, the *in vivo* proteolysis of yBPL observed in bacteria reduces the overall possible yields of recombinant material. By employing protease deficient strains of either *E. coli* or yeast as the host, this problem may be overcome to some extent. An alternative to this approach would be to construct a yBPL analogue where the cleavage motif was altered such that the site became protease insensitive, as discussed in 3.3.4. If this analogue was active and undigested *in vivo*, the gel filtration step could be employed as a final cleanup step rather than a method to separate the uncleaved from the cleaved material. This alone would greatly improve enzyme yields.

8.2 ROLE OF THE N-TERMINAL DOMAIN OF yBPL

Whilst a specific function for the large N-terminal regions of eukaryotic BPLs has not yet been determined, the present study using the yeast enzyme has provided some insights into understanding this region. The protein contains a 26 kDa N-terminal domain connected via a linker to the C-terminal domain containing the biotin and ATP binding determinants. Upon ligand binding, yBPL undergoes conformational changes which renders the protein less sensitive to

proteolysis in both the linker region and substrate binding domain. Deletion of the N-terminal domain generated a molecule which was 3 500-fold less active than the intact enzyme and unable to complement a mutant strain of *E. coli*. It is my view that two approaches will ultimately define function(s) associated with the N-terminal domain. These are *in vivo* studies in yeast and structural determination.

The first approach to defining a role for the yBPL N-terminal domain requires increasing our understanding of biotin metabolism in yeast. We have seen that the molecule central to biotin metabolism in bacteria is BirA. It is plausible that BPL may also play a pivotal role in regulating biotin related events in higher organisms. Before determining the functions of BPL there are several processes relating to biotin metabolism requiring further investigation. For example, the mechanism by which a yeast cell senses cellular concentrations of biotin (and therefore recognises biotin stress) is not known. The level of activity of the biotin transport protein is dependent upon the amount of biotin in the growth media (Rogers & Lichstein, 1969b; Stolz *et al.*, 1999), suggesting that a feedback mechanism exists. How yeast cells respond to biotin stress conditions is also unclear. In the present work, I have shown that conformational changes accompany biotin binding to yBPL. This maybe the structural cue the cell employs to sense the concentration of biotin in the cell and initiate the necessary cellular responses.

A biotin recycling system, analogous to biotinidase in mammals, has not been reported for yeast. These organisms are biotin auxotrophs and, therefore, likely to stringently control their usage of the vitamin. Like BirA, biotinidase is a bifunctional molecule as it is also the serum biotin transport protein (Hymes & Wolf, 1999). Thus, in the biological systems investigated so far, there is a

precedent that the proteins involved in biotin metabolism play multiple roles in biotin-related events.

The second approach available for defining a role for the N-terminal domain is structural determination. X-ray crystallography would be employed as the size of the molecule precludes the use of NMR. In recent years a large number of enzyme structures have appeared in the literature. Structural studies on 'complex enzymes' catalysing more than one reaction have revealed invaluable insights into reaction mechanisms. For example, carbamoyl phosphate synthetase (Holden *et al.*, 1998) and glutamine PRPP amidotransferase (Smith, 1998) contain tunnels running through the molecule which permit channelling of reaction intermediates between the two subsites catalysing the partial reactions. Other studies have investigated the role of allosteric activators and inhibitors upon protein structure. The catalytic cleft of glutamine dehydrogenase is thought to be opened by the binding of ADP and buried by protein structure upon GTP and NADH binding thereby inactivating the enzyme (Peterson & Smith, 1999). Similarly, the N-terminal regulatory domain of phenylalanine hydroxylase behaves like a cap covering the active site and inactivating the enzyme. Upon phenylalanine binding to the regulatory site, conformational changes permit binding of the amino acid to the active site (Vrana, 1999). It is easy to speculate that the N-terminal domain of yBPL may have some role in regulating the activity of the enzyme. Structural determination of yBPL in its various liganded states will reveal whether or not the N-terminal domain does play a role in this regard.

8.3 TRUNCATION ANALYSIS OF yBPL

Chapter 5 described the expression of a series of N-terminally truncated variants of yBPL. These truncations were unable to complement a conditionally lethal *birA*$^-$ strain of *E. coli* grown on selective medium. It is possible that these truncations were incorrectly folded in the bacteria and therefore non-functional. The bacteria posses different protein folding mechanisms which are less suited to the folding of multidomain proteins (Ellis & Hartl, 1999). In order to investigate this further, the analysis of enzyme activity should be repeated using yBPL truncations expressed in yeast.

To determine if the truncations are active *in vivo*, a complementation assay in yeast can be established. No conditional mutants of yBPL have been reported which abolish enzyme activity completely. However Hoja and coworkers (1998) reported a supressor mutant which could be rescued by overexpression of a yeast tRNA gene. However, low levels of BPL activity were detected in this strain suggesting leakiness of the amber mutation. Alternatively, a new yeast strain could be engineered. One of the *bpl* alleles can be disrupted by homologous recombination with a suitable nutritional marker, for example a uracil biosynthetic gene. Upon sporulation, haploid cells with a functional *bpl* gene will survive on non-selective medium and will be selected against upon uracil-deficient medium. This permits the haploid spores with the disrupted *bpl* allele to be used for the complementation assay. On uracil-deficient media, only haploid spores expressing active BPL from an introduced plasmid survive. The N-terminally truncated variants of BPL constructed in this thesis could be analysed by employing this strategy.

8.4 BIOTIN AND ATP BINDING TO yBPL

Upon ATP and biotin binding, BPLs undergo a series of conformational changes. These conformational changes can be viewed as being analogous to steps along a protein folding pathway. Apo BPL can be thought of as a folding intermediate along the pathway to the finally folded product, holo BPL. Thus apo BPL can be equated to a molten globule, as both exhibit extensive secondary structure without having the same conformation as the folded molecule. The conformational changes accompanying ligand binding to BirA have been measured by quenching of the intrinsic tryptophan fluorescence of the protein (Xu & Beckett, 1994). Interestingly, a truncated form of BirA, where the N-terminal domain had been deleted, was shown to bind both biotin and biotinyl-5'-AMP but the ligands did not induce the quenching effect (Xu & Beckett, 1996). This suggests that determinants present in the N-terminal domain of BirA may help catalyse ligand induced protein folding events. Work presented in this thesis showed that deletion of the N-terminal domain of yeast BPL reduced the overall activity of the enzyme. Determinants in the pro-domains of α-lytic protease and subtilisin have also been shown to increase the rate of protein folding reactions (Baker & Agard, 1994). These regions function by allowing stable intermediates in the folding pathway to be converted to the native state. Here I propose that an analogous situation occurs with BPLs. The N-terminal domain, not directly participating in catalysis, is required for the enzyme to fold from the apo form into the active conformation upon ligand binding. Determination of the apo and holo BPL crystal structures will be required to confirm this hypothesis.

8.5 BIOTIN DOMAINS BINDING TO yBPL

Co-crystallisation of BPL with a biotin domain will ultimately reveal the determinants present in both proteins required to specify the protein:protein interaction. However, at present this is a technically difficult process as the BPL : biotinyl-5'-AMP : biotin domain complex must be captured without the biotinyl-transfer reaction proceeding. Currently a non-hydrolysable analogue of biotinyl-5'-AMP is being sought and seems to be the most likely avenue towards success. In the meantime, mutagenesis studies such as those being performed in our laboratory are providing valuable insights into this functional interaction.

Work presented in this thesis has revealed that enzymatic biotinylation of protein substrates requires two key events. Firstly, BPLs recognise and complex with a three dimensional biotin domain. It is clear that the recognition process between BPLs and their protein substrates has a major structural component. Introducing gross conformational changes to an otherwise structured biotin domain abolishes biotinylation. Furthermore, subtle conformational changes to the protein substrate affect the efficiency of the protein:protein interaction. The formation of the holo yBPL-apo biotin domain complex precisely positions the side chain of the biotin accepting lysine into the active site of the enzyme. This brings the reactive ε-amino group into close proximity of the 'activated' form of biotin, biotinyl-5'-AMP, thus allowing the chemical reaction to occur in the active site of the enzyme.

Chemically synthesised biotinyl-5'-AMP reacts readily with molecules containing amine groups. Therefore, in a cellular environment containing an uncountable number of potentially reactive amine groups, one role of a BPL is to prevent non-specific biotinylation. After synthesising biotinyl-5'-AMP, BPLs must

shield the molecule from the cellular environment. Only specific protein substrates, the biotin domains, are biotinylated as only these proteins containing the structural information required to bypass this shield.

The structural determination of a number of enzyme : substrate complexes has shown that the precise positioning of reacting molecules within proteins is essential for enzyme reactions (Davies *et al.*, 1998; Smith, 1998; Leppanen *et al.*, 1999). This permits chemical reactions to proceed between the reactive sites of substrate and the side chains of the enzyme in an isolated environment. Understanding the precise nature by which molecules bind and react with proteins and other macromolecules is already an important field of investigation and has lead to the rational design of inhibitory molecules and potential new drugs. A successful example of this is the work currently being performed on neuraminidase, a surface protein required for the release of influenza virus particles from infected cells (Varghese *et al.*, 1998). The crystal structure of the protein has been determined in both the presence and absence of sialic acid, the natural substrate for the enzyme (Varghese *et al.*, 1998). By identifying residues in the active site of neurominidase which were invariant in all known strains of the virus, sialic acid derived inhibitors which interact specifically with the protein and have K_i values in the nanomolar range have been reported (Varghese *et al.*, 1998). One of these inhibitors has been shown to be effective in animal studies and clinical trials for the treatment of influenza (Varghese *et al.*, 1998). The application of rational drug design to the treatment of human conditions is a powerful and viable technology. Exciting progress has already been achieved for processes such as blocking HIV-1 infection (Michael & Moore, 1999), identifying new anticancer reagents (Hartwell *et al.*, 1997) and designing derivatives of aspirin which deliver better pain relief with fewer side effects (Pennisi, 1998).

Chapter 9

REFERENCES

REFERENCES

Abbott, J. & Beckett, D. (1993). Cooperative binding of the *Escherichia coli* repressor of biotin biosynthesis to the biotin operator sequence. *Biochemistry* **32**(37), 9649-9656.

Al-Feel, W., Chirala, S. S. & Wakil, S. J. (1992). Cloning of the yeast FAS3 gene and primary structure of yeast acetyl CoA carboxylase. *Proc. Natl. Acad. Sci, USA.* **89**, 534-538.

Aoki, Y., Li, X., Sakamoto, O., Hiratsuka, M., Akaishi, H., Xu, L. Q., Briones, P., Suormala, T., Baumgartner, E. R., Suzuki, Y. & Narisawa, K. (1999). Identification and characterization of mutations in patients with holocarboxylase synthetase deficiency. *Hum. Genet.* **104**(2), 143-148.

Aoki, Y., Suzuki, Y., Li, X., Sakamoto, O., Chikaoka, H., Takita, S. & Narisawa, K. (1997). Characterization of mutant holocarboxylase synthetase (HCS) - a *K*m for biotin was not elevated in a patient with HCS deficiency. *Ped. Res.* **42**(6), 849-854.

Aoki, Y., Suzuki, Y., Sakamoto, O., Li, X., Takahashi, K., Ohtake, A., Sakuta, R., Ohura, T., Miyabayashi, S. & Narisawa, K. (1995). Molecular analysis of holocarboxylase synthetase deficiency - a missense mutation and a single base deletion are predominant in Japanese patients. *Biochim. Biophys. Acta. Mol. Bas. Dis.* **1272**(3), 168-174.

Artymiuk, P. J., Rice, D. W., Poirrette, A. R. & Willet, P. (1994). A tale of two synthetases. *Nat. Struct. Biol.* **1**(11), 758-760.

Athappilly, F. K. & Hendrickson, W. A. (1995). Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by Mad phasing. *Structure* **3**(12), 1407-1419.

Attwood, P. V. (1995). The structure and mechanism of action of pyruvate carboxylase. *Int. J. Biochem. Cell. Biol.* **27**, 231-249.

Attwood, P. V., Johannssen, W., Chapman-Smith, A. & Wallace, J. C. (1993). The existance of multiple tetrameric conformers of chicken liver pyruvate carboxylase and their roles in dilution inactivation. *Biochem. J.* **290**, 583-590.

Atwell, S. & Wells, J. A. (1999). Selection for improved subtiligases by phage display. *Proc. Natl. Acad. Sci. USA* **96**(17), 9497-9502.

Baker, D. & Agard, A. (1994). Kinetics versus thermodynamics in protein folding. *Biochemistry* **33**, 7505-7509.

Barker, D. F. & Campbell, A. M. (1981a). The *birA* gene of *Escherichia coli* encodes a biotin holoenzyme synthetase. *J. Mol. Biol.* **146**, 451-467.

Barker, D. F. & Campbell, A. M. (1981b). Genetic and biochemical characterisation of the *birA* gene and its product : Evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in *Escherichia coli. J. Mol. Biol.* **146**, 469-492.

Bass, S., Greene, R. & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins* **8**, 309-314.

Beckett, D., Kovaleva, E. & Schatz, P. J. (1999). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Prot. Sci.* **8**(4), 921-929.

Beckett, D. & Matthews, B. W. (1997). *Escherichia coli* repressor of biotin biosynthesis. *Methods Enzymol.* **279**, 362-376.

Berg, A., Vervoort, J. & Dekok, A. (1996). Solution structure of the lipoyl domain of the 2-oxyglutarate dehydrogenase complex from *Azotobacter vinelandii*. *J. Mol. Biol.* **261**(3), 432-442.

Berg, A., Vervoort, J. & Dekok, A. (1997). Three-dimensional structure in solution of the N-terminal lipoyl domain of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **244**(2), 352-360.

Best, E. A. & Knauf, V. C. (1993). Organisation and nucleotide sequences of the genes encoding the biotin carboxyl carrier protein and biotin carboxylase protein of *Pseudomonas aeroginosa* acetyl coenzyme A carboxylase. *J. Bacteriol.* **175**, 6881-6889.

Blom, W. & Scholte, H. R. (1981). Acetyl CoA carboxylase deficiency: An inborn error of *de novo* fatty acid synthesis. *N. Eng. J. Med.* **305**, 465.

Bordo, D. & Argos, P. (1991). Suggestions for safe residue substitutions in site directed mutagenesis. *J. Mol. Biol.* **217**, 721-729.

Bower, S., Perkins, J., Yocum, R. R., Serror, P., Sorokin, A., Rahaim, P., Howitt, C. L., Prasad, N., Ehrlich, S. D. & Pero, J. (1995). Cloning and characterization of the *Bacillus subtilis* BirA gene encoding a repressor of the biotin operon. *J. Bacteriol.* **177**(9), 2572-2575.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.

Brewster, N. K., Val, D. L., Walker, M. E. & Wallace, J. C. (1994). Regulation of pyruvate carboxylase isozyme (Pyc1, Pyc2) gene expression in

Saccharomyces cerevisiae during fermentative and nonfermentative growth. *Arch. Biochem. Biophys.* **311**(1), 62-71.

Brocklehurst, S. M. & Perham, R. N. (1993). Prediction of the three dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: A new automated method for the prediction of protein tertiary structure. *Protein Sci.* **2**, 626-639.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). Charmm: A programme for macromolecular energy minimalisation and dynamics calculations. *J. Comput. Chem.* **4**, 187-217.

Brosius, J. & Holy, A. (1984). Regulation of ribosomal RNA promotors with a synthetic lac operator. *Proc. Natl. Acad. Sci. USA* **81**, 6929-6933.

Browner, M. F., Taroni, F., Sztul, E. & Rosenberg, L. E. (1989). Sequence analysis, biogenesis and mitochondrial import of the α-subunit of rat liver propionyl CoA carboxlase. *J. Biol. Chem.* **264**, 12680-12685.

Brunger, A. T. (1992). X-PLOR Version 3.1: A system for X-ray crystallography, Yale University, New Haven, CT.

Buchli, P. J., Wu, Z. & Ciardelli, T. L. (1997). The functional display of interleukin-2 on filamentous phage. *Arch. Biochem. Biophys.* **339**, 79-84.

Bult, C. J. *et al.* (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii. Science* **273**(5278), 1058-1073.

Buoncristiani, M. R. & Otsuka, A. J. (1988). Overproduction and rapid purification of the biotin operon repressor from *Escherichia coli.J. Biol. Chem.* **263**, 1013-1016.

Burri, B. J., Sweetman, L. & Nyhan, W. L. (1981). Mutant holocarboxylase synthetase. *J. Clin. Invest.* **68**, 1491-1495.

Cagnon, C., Valverde, V. & Masson, J.-M. (1991). A new family of sugar-inducible expression vectors for *Escherichia coli.Prot. Eng.* **4**, 843-847.

Caparon, M. H., Ciechi, P. A., Devine, C. S., Olins, P. O. & Lee, S. C. (1995). Analysis of novel streptavidin-binding peptides, identified using a phage display library, shows that amino acids external to a perfectly conserved consensus sequence and to the presented peptides contribute to binding. *Mol. Divers.* **1**, 241-246.

Caspari, T., Robl, I., Stolz, J. & Tanner, W. (1996). Purification of the Chlorella hup1 hexose-proton symporter to homogeneity and its reconstitution *in vitro*. *Plant J.* **10**(6), 1045-1053.

Cazzulo, J. J. & Stopani, A. O. M. (1968). The regulation of yeast pyruvate carboxylase by acetyl-coenzyme A and L-aspartate. *Arch. Biochem. Biophys.* **127**, 563-567.

Cazzulo, J. J., Sundaram, T. K., Dilks, S. N. & Kornberg, H. L. (1971). Synthesis of pyruvate carboxylase from its apoenzyme and (+)-biotin in *Bacillus stearothermophilus*. Purification and properties of the apoenzyme and the holoenzyme synthetase.*Biochem. J.*(122), 653-661.

Cesareni, G. (1992). Peptide display on filamentous phage capsids: a new powerful tool to study protein-ligand interaction. *FEBS* **307**(1), 66-70.

Chang, H. I. & Cohen, N. D. (1983). Regulation and intracellular localisation of the biotin holocarboxylase synthetase of 3T3-L1 cells. *Arch. Biochem. Biophys.* **225**, 237-347.

Chapman-Smith, A. & Cronan, J. E. (1999a). Molecular biology of biotin attachment to proteins. *J. Nutr.* **129**(2 Suppl S), 477s-484s.

Chapman-Smith, A. & Cronan, J. E. J. (1999b). The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biol. Sci.* **24**(9): 359-363

Chapman-Smith, A., Forbes, B. E., Wallace, J. C. & Cronan, J. E. (1997). Covalent modification of an exposed surface turn alters the global conformation of the biotin carrier domain of *Escherichia coli* acetyl-CoA carboxylase. *J. Biol. Chem.* **272**(41), 26017-26022.

Chapman-Smith, A., Morris, T. W., Wallace, J. C. & Cronan, J. E. (1999). Molecular recognition in a post-translational modification of exceptional specificity - Mutants of the biotinylated domain of acetyl-CoA carboxylase defective in recognition by biotin protein ligase. *J. Biol. Chem.* **274**(3), 1449-1457.

Chapman-Smith, A., Turner, D. L., Cronan, J. E., Morris, T. W. & Wallace, J. C. (1994). Expression, biotinylation and purification of a biotin-domain peptide from the biotin carboxy carrier protein of *Escherichia coli* acetyl-CoA carboxylase. *Biochem. J.* **302**(3), 881-887.

Chiba, Y., Suzuki, Y., Aoki, Y., Ishida, Y. & Narisawa, K. (1994). Purification and properties of bovine liver holocarboxylase synthetase. *Arch. Biochem. Biophys.* **313**(1), 8-14.

Christner, J. E., Schlesimger, M. & Coon, M. J. (1964). Enzymatic activation of biotin. *J Biol. Chem.* **239**, 3997-4005.

Clackson, T. & Wells, J. A. (1994). *In vitro* selection from protein and peptide libraries. *Trends Biotech.* **12**(5), 173-184.

Cleland, W. W. (1970). Steady State Kinetics. *The Enzymes* **2**, 1-65

Cohen, N. D., Duc, J. A., Beegan, H. & Utter, M. F. (1979). Quaternary structure of pyruvate carboxylase from *Pseudomonas citronellolis*. *J. Biol. Chem.* **254**, 9262-9269.

Cohen, N. D., Thomas, M. & Stack, M., Eds. (1985). Biotin. Vol. 447. Ann. New York. Acad. Sci. Edited by Dakshinamurti, K. & Bhagavan, H. N.

Cooper, T. G. & Wood, H. G. (1971). The carboxylation of phosphoenolpyruvate and pyruvate. II. The active species of "CO2" utilized by phosphoenlpyruvate carboxylase and pyruvate carboxylase.*J. Biol. Chem.* **246**, 5488-5490.

Corey, D. R., Shiau, A. K., Yang, Q., Janowski, B. A. & Craik, C. S. (1993). Trypsin display on the surface of bacteriophage. *Gene* **128**, 129-134.

Crissman, J. W. & Smith, G. P. (1984). Gene-III protein of filamentous phages: evidence for a carboxyl-terminal domain with a role in morphogenesis. *Virology* **132**, 445-455.

Cronan, J. E., Jnr. (1990). Biotination of proteins *in vivo*. *J. Biol. Chem.* **265**, 10327-10333.

Cronan, J. E. J. & Wallace, J. C. (1995). The gene encoding the biotin-apoprotein ligase of *Saccharomyces cerevisiae*. *FEMS Microbiol. Letts* **130**(2-3), 221-229.

Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990). Peptides on phage: A vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. USA* **87**, 6378-6382.

Dahmane, N., Ghezala, G. A., Gosset, P., Chamoun, Z., Dufresnezacharia, M. C., Lopes, C., Rabatel, N., Gassanovamaugenre, S., Chettouh, Z., Abramowski, V., Fayet, E., Yaspo, M. L., Korn, B., Blouin, J. L., Lehrach, H., Poutska, A., Antonarakis, S. E., Sinet, P. M., Creau, N. & Delabar, J. M.

(1998). Transcriptional Map of the 2.5-Mb Cbr-Erg Region of Chromosome 21 Involved in Down-Syndrome. *Genomics* **48**(1), 12-23.

Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993). Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.

Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schulein, M. & Withers, S. G. (1998). Snapshots along an enzymatic reaction coordinate: Analysis of a retaining β-glycoside hydrolase. *Biochemistry* **37**, 11707-11713.

Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L., Neri, P., Winter, G. & Neri, D. (1999). A strategy for the isolation of catalytic activities from repertoires of enzymes displayed on phage. *J. Mol. Biol.* **286**(2), 617-633.

Dimroth, P. (1985). Biotin-dependent decarboxylases as energy transducing systems.*Ann. N. Y. Acad. Sci.* **447**, 72-85.

Dombrosky, P. M., Schmid, M. B. & Young, K. D. (1994). Sequence divergence of the Murb and Rrfb genes from *Escherichia coli* and *Salmonella typhimurium*. *Arch. Microbiol.* **161**(6), 501-507.

Downer, W. J., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation by high voltage electroporation. *Nucl. Acids Res.* **16**(13), 6127-6145.

du Vigneaud, V., Melville, D. B., Folkers, K., Wolf, D. E. & Mozingo, R. (1942). The structure of biotin: A study of desthiobiotin. *J. Biol. Chem.* **146**, 475-85.

du Vigneaud, V., Melville, D. B., Gyorgy, P. & Rose, C. S. (1940). On the identity of vitamin H with biotin. *Science* **92**, 62-63.

Duffy, S., Tsao, K. L. & Waugh, D. S. (1998). Site-specific, enzymatic biotinylation of recombinant proteins in *Spodoptera frugiperda* cells using biotin acceptor peptides. *Anal. Biochem.* **262**(2), 122-128.

Dunn, I. S. (1996). Phage display of proteins. *Curr. Opin. Biotech.* **7**, 547-553.

Dupuis, L., Leondelrio, A., Leclerc, D., Campeau, E., Sweetman, L., Saudubray, J. M., Herman, G., Gibson, K. M. & Gravel, R. A. (1996). Clustering of mutations in the biotin-binding region of holocarboxylase synthetase in biotin-responsive multiple carboxylase deficiency. *Hum. Mol. Gen.* **5**(7), 1011-1016.

Dyer, D. L. & Said, H. M. (1997). Biotin uptake in cultured cell lines. *Methods Enzmol.* **279**, 393-405.

Eakin, R. E., Snell, E. E. & Williams, R. J. (1941). *J. Biol. Chem.* **140**, 535.

Eisenberg, M. A. (1973). Biotin: biogenesis, transport, and their regulation. *Adv Enzymol. Relat. Areas Mol. Biol.* 38:317-72.

Eisenberg, M. A., Prahash, O. & Hsiung, S. C. (1982). Purification and properties of the biotin repressor. A bifunctional protein. *J. Biol. Chem.* **257**, 15167-15173.

Ellis, R. J. & Hartl, F. U. (1999). Principles of protein folding in the cellular environment. *Curr. Opin. Struct. Biol.* **9**, 102-110.

Fersht, A. (1985). In *Enzyme Structure and Function*, pp. 212-214. W. H. Freeman, New York.

Fisher, D. & Eisenberg, D. (1996). Protein fold recognition using sequence-derived predictions. *Protein Sci.* **5**, 947-955.

Freytag, S. O. & Collier, K. J. (1984). Molecular cloning of a cDNA for human pyruvate carboxylase: Structural relationship to other biotin-containing

carboxylases and regulation of mRNA content in differentiating preadipocytes. *J. Biol. Chem.* **259**, 12831-12837.

Friedrich, W. (1988). Biotin. In *Vitamins*, pp. 753-805. de Gruyter, New York.

Friesen, W. J. & Darby, M. K. (1997). Phage display of RNA binding zinc fingers from transcription factor IIIA. *J. Biol. Chem.* **272**(17), 10994-10997.

Fuchs, J., Johannssen, W., Rhode, M. & Mayer, F. (1988). Pyruvate carboxylase from *Pseudomonas citronellolis*: Shape of the enzyme and localisation of its prosthetic group by electron microscope affinity labelling. *FEBS Letts.* **231**, 102-106.

Genbauffe, F. S. & Cooper, T. G. (1986). Induction and repression of the urea amidolyase gene in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **6**, 3954-3964.

Genbauffe, F. S. & Cooper, T. G. (1991). The urea amidolyase (DUR1,2) gene of *Saccharomyces cerevisiae*. *DNA Seq.* **2**, 19-32.

Geschwind, D. H., Rhee, R. & Nelson, S. F. (1996). A biotinylated MutS fusion protein and its use in a rapid muation screening technique. *Genet. Anal.* **13**, 105-111.

Gheber, L., Kuo, S. C. & Hoyt, M. A. (1999). Motile properties of the kinesin-related Cin8p spindle motor extracted from *Saccharomyces cerevisiae* cells. *J. Biol. Chem.* **274**(14): 9564-72

Goss, J. A., Cohen, N. D. & Utter, M. F. (1981). Characterisation of the subunit structure of pyruvate carboxylase from *Pseudomonas citronellolis*. *J. Biol. Chem.* **256**, 11819-11825.

Goss, N. H., Dyer, P. Y., Keech, D. B. & Wallace, J. C. (1979). An electron microscopic study of pyruvate carboxylase. *J. Biol. Chem.* **254**; 1734-1739.

Gram, H., Strittmatter, U., Lorenz, M., Gluck, D. & Zenke, G. (1993). Phage display as a rapid gene expression system: Production of bioactive cytokine-phage and generation of neutralizing monoclonal antibodies. *J. Immunol. Methods* **161**, 169-176.

Gray, C. W., Brown, R. S. & Marvin, D. A. (1981). Adsorption complex of filamentous fd virus. *J. Mol. Biol.* **146**, 621-627.

Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995). Three dimensional structure of a lipoyl domain from the dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.

Green, N. M. (1975). Avidin. *Adv. Prot. Chem.* **29**, 85-113.

Greenwood, J., Wallis, A. E. & Perham, R. N. (1991). Multiple display of foreign peptides on a filamentous bacteriophage. *J. Mol. Biol.* **220**, 821-827.

Gu, H., Yi, Q., Bray, S. T., Riddle, D. S., Shiau, A. K. & Baker, D. (1995). A phage display system for studying the sequence determinants of protein folding. *Protein Sci.* **4**, 1108-1117.

Gyorgy, P., Rose, C. S., Hofmann, K., Melville, D. B. & du Vigneaud, V. (1940). A further note on the identity of vitamin H with biotin. *Science* **92**, 609.

Ha, J., Daniel, S., Kong, I.-S., Park, C.-K., Tae, H.-J. & Kim, K.-H. (1994). Cloning of the human acetyl CoA carboxylase cDNA. *Eur. J. Biochem.* **219**, 297-306.

Hablacher, M., Ivessa, A. S., Paltauf, F. & Kohlwein, S. D. (1993). Acetyl-CoA carboxylase from yeast is an essential enzyme and is regulated by factors that control phospholipid metabolism. *J. Biol. Chem.* **268**, 10946-10952.

Harris, S. A., Wolf, D. E., Mozingo, R. & Folkers, K. (1943). Synthetic biotin. *Science* **97**, 447-48.

Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. (1997). Integrating genetic approaches into the discovery of anticancer drugs. *Science* **278**, 1064-1068.

Hiratsuka, M., Sakamoto, O., Li, X., Suzuki, Y., Aoki, Y. & Narisawa, K. (1998). Identification of holocarboxylase synthetase (HCS) proteins in human placenta. *Biochim. Biophys. Acta Prot. Struct. Mol. Enzymol.* **1385**(1), 165-171.

Hoja, U., Wellein, C., Greiner, E. & Schweizer, E. (1998). Pleiotropic phenotype of acetyl-CoA carboxylase defective yeast cells. *Eur. J. Biochem.* **254**, 520-526.

Holden, H. M., Thoden, J. B. & Raushel, F. M. (1998). Carbamoyl phosphate synthetase: A tunnel runs through it. *Curr. Opin. Struct. Biol.* **8**, 679-685.

Holliger, P. & Riechmann, L. (1997). A conserved infection pathway for filamentous bacteriophages is suggested by the structure of the membrane penetration domain of the minor coat protein g3p from phage fd. *Structure* **5**, 265-275.

Holmes, D. I., Manafounis, D., Ward, W. H. J. & Leatherbarrow, R. J. (1996). Bacteriophage display of chymotrypsin inhibitor 2. *Prot. Pept. Letts.* **6**, 415-422.

Hoogenboom, H. R., Griffiths, D., Johnson, K. S., Chiswell, D. J., Hudson, P. & Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucl. Acids Res.* **19**(15), 4133-4137.

Howard, P. K., Shaw, J. & Otsuka, A. J. (1985). Nucleotide sequence of the *birA* gene encoding the biotin operon repressor and biotin holoenzyme synthetase functions of *Escherichia coli*. *Gene* **35**, 321-331.

Hudson, P. J. (1998). Recombinant antibody fragments. *Curr. Opin. Biotech.* **9**(4), 395-402.

Hymes, J. & Wolf, B. (1999). Human biotinidase isn't just for recycling biotin. *J. Nutr.* **129**, 485S-489S.

Irias, J. J., Olmsted, M. R. & Utter, M. F. (1969). Pyruvate carboxylase: Reversible inactivation by cold. *Biochemistry* **8**, 5136-5148.

Jander, G., Cronan, J. E. & Beckwith, J. (1996). Biotinylation *in vivo* as a sensitive indicator of protein secretion and membrane protein insertion. *J. Bacteriol.* **178**(11), 3049-3058.

Jitrapakdee, S. (1999). Characterisation of the pyruvate carboxylase gene and studies on the regulation of its expression in rat, PhD Thesis, University of Adelaide.

Jitrapakdee, S., Booker, G. W., Cassady, A. I. & Wallace, J. C. (1996). Cloning, sequencing and expression of rat liver pyruvate carboxylase. *Biochem. J.* **316**, 631-637.

Jitrapakdee, S. & Wallace, J. C. (1999). Structure, function and regulation of pyruvate carboxylase. *Biochem. J.* **340**(1), 1-16.

Johannssen, W., Attwood, P. V., Wallace, J. C. & Keech, D. B. (1983). Localisation of the active site of pyruvate carboxylase by electron microscopic examination of avidin-enzyme complexes. *Eur. J. Biochem* **133**, 201-206.

Kang, A. S., Barbas, C. F., Janda, K. D., Benkovic, S. J. & Lerner, R. A. (1991). Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA* **88**, 4363-4366.

Kaziro, Y., Hass, L. F., Boyer, P. D. & Ochoa, S. (1962). Mechanism of the propionyl carboxylase reaction II. Isotopic exchange and tracer experiments *J. Biol. Chem.* **237**, 1460- 68.

Keech, D. B. & Utter, M. F. (1963). Pyruvate carboxylase II: Properties. *J. Biol. Chem.* **238**, 2609-2614.

Khew-Goodall, Y. S., Johannssen, W., Attwood, P. V. & Wallace, J. C. (1991). Studies on dilution inactivation of sheep liver pyruvate carboxylase. *Arch. Biochem. Biophys.* **284**, 98-105.

Klenk, H. P. *et al.* (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**(6658), 364-370.

Knowles, J. R. (1989). The mechanism of biotin-dependent enzymes. *Ann. Rev. Biochem.* **58**, 195-221.

Kogl, F. & Tonnis, B. (1936). *Z. Physiol. Chem.* **242**, 43.

Kondo, H., Uno, S., Komizo, Y. Y. & Sunamoto, J. (1984). 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. Peptide Protein Res.* **23**, 559-564.

Koradi, R., Billeter, M. & Wuthrich, K. (1996). Molmol - A program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**(1), 51-55.

Korpela, J. (1984). Avidin, a high affinity biotin-binding protein as a tool and subject of biological research. *Med. Biol.* **62**, 62-65.

Kosow, D. P., Huang, S. C. & Lane, M. D. (1962). Propionyl holocarboxylase synthesis. *J. Biol. Chem.* **237**, 3633-3639.

Kraulis, P. J. (1991). Molscript: A programme to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lamhonwah, A.-M., Quan, F. & Gravel, R. A. (1987). Sequence homology around the biotin-binding site of human propionyl CoA carboxylase and pyruvate carboxylase. *Arch. Biochem. Biophys.* **254**, 631-636.

Lamhonwah, A. M., Barankiewicz, T. J., Willard, H. F., Mahuran, D. J., Quan, F. & Gravel, R. A. (1986). Isolation of cDNA clones coding for the alpha and beta chains of human propionyl CoA carboxylase: chromosomal assignments and DNA polymorphisms associated with PCCA and PCCB genes. *Proc. Natl. Acad. Sci. USA* **83**, 4864-4868.

Lane, D. M., Young, D. L. & Lynen, F. (1964). The enzymatic synthesis of holocarboxylase from apotranscarboxylase and (+) biotin. *J. Biol. Chem.* **239**, 2858-2864.

Lane, M. D., Moss, J. D. & Polakis, S. E. (1974). Acetyl coenzyme A carboxylase. *Curr. Top. Cell regul.* **8**, 139-195.

Lee, H. J. & Wilson, I. B. (1971). On the structure of renin substrate *Biochim. Biophys. Acta* **242**, 519-522.

Leon-Del-Rio, A. & Gravel, R. A. (1994). Sequence requirements for the biotinylation of carboxyl-terminal fragments of human propionyl-CoA carboxylase alpha subunit expressed in *Escherichia coli*. *J. Biol. Chem.* **269**(37), 22964-22968.

Leon-Del-Rio, A., Leclerc, D., Akerman, B., Wakamatsu, N. & Gravel, R. A. (1995). Isolation of a cDNA encoding human holocarboxylase synthetase by functional complementation of a biotin auxotroph of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **92**(10), 4626-4630.

Leppanen, V.-M., Merckel, M. C., Ollis, D. L., Wong, K. K., Kozarich, J. W. & Goldman, A. (1999). Pyruvate formate lyase is structurally homologous to type 1 ribonucleotide reductase. *Structure* **7**, 733-744.

Lerner, C. G. & Saiki, A. Y. C. (1996). Scintillation proximity assay for human DNA topoisomerase I using recombinant biotinyl-fusion protein produced in baculovirus-infected insect cells. *Anal. Biochem.* **240**, 185-196.

Li, S. & Cronan, J. E. J. (1992). The gene encoding the biotin carboxylase subunit of *Escherichia coli* acetyl Co-A carboxylase. *J. Biol. Chem.* **267**, 855-863.

Lien, S. (2000). PhD Thesis, University of Adelaide.

Lim, F., Morris, C. P., Occhiodoro, F. & Wallace, J. C. (1988). Sequence and domain structure of yeast pyruvate carboxylase. *J. Biol. Chem.* **263**(23), 11493-11497.

Lim, F., Rhode, M., Morris, C. P. & Wallace, J. C. (1987). Pyruvate carboxylase in the yeast *pyc* mutant. *Arch. Biochem. Biophys.* **258**, 259-264.

Lowman, H. B. (1997). Bacteriophage display and discovery of peptide leads for drug development. *Ann. Rev. Biophys. Biomol. Struct.* **26**, 401-24.

Lowman, H. B. & Wells, J. A. (1991). Monovalent phage display: a method for selecting variant proteins from random libraries. *Methods: Companion Methods Enzymol.* **3**(3), 205-216.

Lubkowski, J., Hennecke, F., Pluckthun, A. & Wlodawer, A. (1998). The structural basis of phage display elucidated by the crystal structure of the N-terminal domains of g3p. *Nat. Struct. Biol.* **5**, 140-147.

Lucic, M. R., Forbes, B. E., Grosvenor, S. E., Carr, J. M., Wallace, J. C. & Forsberg, G. (1998). Secretion in *Escherichia coli* and phage-display of recombinant insulin-like growth factor binding protein-2. *J. Biotech.* **61**(2), 95-108.

Lynen, F. (1975) In *Energy regulation and biosynthesis in molecular biology* (Richter, D., ed.), pp. 671-698. Walter de Gruyter, Berlin.

Lynen, F., Knappe, J., Lorch, E., Jutting, G. & Ringelmann, E. (1959). *Angew. Chem.* **71**(481).

Madappally, M. M. & Mistry, S. P. (1970). Synthesis of chicken liver pyruvate carboxylase *in vivo* and *in vitroBiochim. Biophys. Acta* **215**, 316-322.

Makowski, L. (1994). Phage display: structure, assembly and engineering of filamentous bacteriophage M13. *Curr. Opin. Struct. Biol.* **4**, 225-230.

Maloy, W. L., Bowien, B. U., Zwolinski, G. K., Kumar, K. G., Wood, H. G., Ericsson, L. H. & Walsh, K. A. (1979). Amino acid sequence of the biotinyl subunit from transcarboxylase. *J. Biol. Chem.* **254**, 11615-11622.

Marvin, D. A. (1998). Filamentous phage structure, infection and assembly. *Curr. Opin. Struct. Biol.* **8**(2), 150-158.

Matsudaira, P. (1988). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes*J. Biol. Chem.* **262**, 10035-10038.

Matthews, D. J., Goodman, L. J., Gorman, C. M. & Wells, J. A. (1994). A survey of furin substrate specificity using substrate phage display. *Prot. Sci.* **3**, 1197-1205.

Matthews, D. J. & Wells, J. A. (1993). Substrate phage: Selection of protease substrates by monovalent phage display. *Science* **260**, 1113-1117.

Mayer, F., Wallace, J. C. & Keech, D. B. (1980). Further electron microscope studies on pyruvate carboxylase. *Eur. J. Biochem.* **112**, 265-272.

McAllister, H. C. & Coon, M. J. (1966). Further studies on the properties of liver propionyl coenzyme A holoenzyme synthetase and the specificity of holocarboxylase formation. *J. Biol. Chem.* **241**, 2855-2861.

McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains *Nature* **348**(6301), 552-554.

McCafferty, J., Jackson, R. H. & Chiswell, D. J. (1991). Phage display expression and affinity chromatography of functional alkaline phosphatase on the surface of bacteriophage. *Prot. Eng.* **4**, 955-961.

McConnell, S. J. & Hoess, R. H. (1995). Tendamistat as a scaffold for comformationally constrained phage peptide libraries. *J. Mol. Biol.* **250**, 460-470.

Merritt, E. A. & Bacon, D. J. (1997). Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Michael, N. L. & Moore, J. P. (1999). HIV-1 entry inhibitors: Evading the tissue. *Nat. Med.* **5**(7), 740-741.

Minich, W. B., Weymayer, J. D. & Loos, U. (1998). Expression of a biotinylated human thyrotropin receptor in HeLa cells using recombinant vaccinia virus and its application for the detection of Graves' autoantibodies. *Thyroid* **8**, 3-7.

Mishina, M., Roggenkamp, R. & Schweizer, E. (1980). Yeast mutants defective in acetyl coenzyme A carboxylase and biotin : apocarboxylase ligase. *Eur. J. Biochem.* **111**, 79-87.

Morita, J., Thuy, L. P. & Sweetman, L. (1998). Deficiency of biotinyl-AMP synthetase activity in fibroblasts of patients with holocarboxylase synthetase deficiency. *Mol. Genet. Metabol.* **64**(4), 250-255.

Mukhopadhyay, B., Stoddard, S. F. & Wolfe, R. (1998). Purification, regulation and molecular and biochemical characertisation of pyruvate carboxylase

from *Methanobacterium thermoautotrophicum* strain delta H. *J. Biol. Chem.* **273**, 5155-5156.

Munkres, K. D. & Richards, F. M. (1965). *Arch. Biochem. Biophys.* **109**, 466-478.

Murthy, P. N. A. & Mistry, S. P. (1974). *In vitro* synthesis of propionyl-CoA holocarboxylase by a partially purified mitochondrial preparation from biotin deficient chicken liver *Can. J. Biochem.* **52**, 800-803.

Murtif, V. L., Bahler, C. R. & Samols, D. (1985). Cloning and expression of the 1.3S biotin-containing subunit of transcarboxylase. *Proc. Natl. Acad. Sci. USA* **82**, 5617-5621.

Murtif, V. L. & Samols, D. (1987). Mutagenesis affecting the carboxyl terminus of the biotinyl subunit of transcarboxylase. *J. Biol. Chem.* **262**, 11813-11816.

Nenortas, E. & Beckett, D. (1996). Purification and characterization of intact and truncated forms of the *Escherichia coli* biotin carboxyl carrier subunit of acetyl-CoA carboxylase. *J. Biol. Chem.* **271**(13), 7559-7567.

Netzer, W. J. & Hartl, F. U. (1997). Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388**, 343-349.

Noble, M. E. M., Musacchio, A., Saraste, M., Courtneidge, S. A. & Wierenga, R. K. (1993). 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**(7), 2617-2624.

O'Neil, K. T. & Hoess, R. H. (1995). Phage display - Protein engineering by directed evolution. *Curr. Opin. Struct. Biol.* **5**(4), 443-449.

Oliver, D. B., Cabelli, R. J., Dolan, K. M. & Jarosik, G. P. (1990). Azide resistant mutants of *Escherichia coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. *Proc. Natl. Acad. Sci. USA* **97**, 8227-8231.

Osmani, S. A., Mayer, F., Marston, F. A. O., Selmes, I. P. & Scrutton, M. C. (1984). Pyruvate carboxylase from *Aspergillus nidulans*: Effects of regulatory modifiers on the structure of the enzyme. *Eur. J. Biochem.* **139**, 509-518.

Parsons, H. T., Lease, J. G. & Kelly, E. (1937). The interrelationship between dietry egg white and requirement for a protective factor in the cure of the nutritive disorder due to egg white. *Biochem J.* **31**, 424-433.

Pennisi, E. (1998). Building a better aspirin. *Science* **280**, 1191-1192.

Peterson, P. E. & Smith, T. J. (1999). The structure of bovine glutamate dehydrogenase provides insights into the mechanism of allostery. *Structure* **7**, 769-782.

Phalip, V., Kuhn, I., Lemoine, Y. & Jeltsch, J. M. (1999). Characterization of the biotin biosynthesis pathway in *Saccharomyces cerevisiae* and evidence for a cluster containing BI05, a novel gene involved in vitamer uptake. *Gene* **232**(1), 43-51.

Phillips, N. F. B., Snoswell, M. A., Chapman-Smith, A., Keech, D. B. & Wallace, J. C. (1992). Isolation of a carboxylphosphate intermediate and the locus of acteyl-CoA action in the pyruvate carboxylase reaction. *Biochemistry* **31**, 9445-9450.

Piffeteau, A. & Gaudry, M. (1985). Biotin uptake: influx, efflux and countertransport in *Escherichia coli* K12 *Biochim. Biophys. Acta* **816**, 77-82.

Polyak, S. W., Forsberg, G., Forbes, B. E., McNeil, K. A., Aplin, S. & Wallace, J. C. (1997). Introduction of spacer peptides N-terminal to a cleavage recognition motif in recombinant fusion proteins can improve for site specific cleavage. *Protein Eng.* **10**(6), 615-619.

Pos, K. M., Bott, M. & Dimroth, P. (1994). Purification of two active fusion proteins of the Na(+)-dependent citrate carrier of *Klebsiella pueumoniae. FEBS Letts* **247**(37-41).

Rebar, E. J. & Pabo, C. O. (1994). Zinc finger phage: Affinity selection of fingers with new DNA binding specificities. *Science* **263**, 671-673.

Reche, P., Li, Y. L., Fuller, C., Eichhorn, K. & Perham, R. N. (1998). Selectivity of post-translational modification in biotinylated proteins - the carboxy carrier protein of the acetyl-CoA carboxylase of *Escherichia coli. Biochem. J.* **329**(3), 589-596.

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

Reddy, D. V., Rothemund, S., Shenoy, B. C., Carey, P. R. & Sonnichsen, F. D. (1998). Structural characterization of the entire 1.3S subunit of transcarboxylase from *Propionibacterium shermanii. Protein Sci.* **7**(10), 2156-2163.

Reddy, D. V., Shenoy, B. C., Carey, P. R. & Sonnichsen, F. D. (1997). Absence of observable biotin-protein interactions in the 1.3S subunit of transcarboxylase: An NMR study. *Biochemistry* **36**, 14676-14682.

Rhode, M., Lim, F. & Wallace, J. C. (1986). Pyruvate carboxylase from *Saccharomyces cerevisiae*: Quaternary structure, effects of allosteric ligands and binding of avidin. *Eur. J. Biochem.* **156**, 15-22.

Rhode, M., Lim, F. & Wallace, J. C. (1991). Electron microscope localisation of pyruvate carboxylase in rat liver and *Saccharomyces cerevisiae* by immunogold proceedures. *Arch. Biochem. Biophys.* **290**, 197-201.

Riechmann, L. & Holliger, P. (1997). The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*. *Cell* **90**, 351-360.

Roberts, B. L., Markland, W., Ley, A. C., Kent, R. B., White, D. W., Guterman, S. K. & Lander, C. (1992). Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage*Proc. Natl. Acad. Sci. USA* **89**, 2429-2433.

Roberts, E. L., Shu, N. C., Howard, M. J., Broadhurst, R. W., Chapman-Smith, A., Wallace, J. C., Morris, T., Cronan, J. E. & Perham, R. N. (1999). Solution structures of apo and holo biotinyl domains from acetyl coenzyme A carboxylase of *Escherichia coli* determined by triple-resonance nuclear magnetic resonance spectroscopy. *Biochemistry* **38**(16), 5045-5053.

Rogers, T. O. & Lichstein, H. C. (1969a). Characterisation of the biotin transport system in *Saccharomyces cerevisiae*. *J. Bacteriol.* **100**(2), 557-564.

Rogers, T. O. & Lichstein, H. C. (1969b). Regulation of biotin transport in *Saccharomyces cerevisiae*. *J. Bacteriol.* **100**(2), 565-572.

Roggenkamp, R., Numa, S. & Schweizer, E. (1980). Fatty acid-requiring mutant *Saccharomyces cerevisiae* defective in acetyl-CoA carboxylase. *Proc. Natl. Acad. Sci. USA.* **77**, 1814-1817.

Rost, B. & Saunder, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* **232**, 584-599.

Russell, R. B. & Barton, G. J. (1993). An SH2-SH3 domain hybrid. *Nature* **364**, 765.

Rylatt, D., B,, Keech, D. B. & Wallace, J. C. (1977). Pyruvate carboxylase: Isolation of biotin containing tryptic peptide and determination of its primary sequence. *Arch. Biochem. Biophys.* **183**, 113-122.

Saggio, I. & Laufer, R. (1993). Biotin binders selected from a random peptide library expressed on phage. *Biochem. J.* **293**, 613-616.

Said, H. M. (1999). Cellular uptake of biotin: Mechanisms and regulation. *J. Nutr.* **129**, 490S-493S.

Sakamoto, O., Suzuki, Y., Aoki, Y., Li, X., Hiratsuka, M., Yanagihara, K., Inui, K., Okabe, T., Yamaguchi, S., Kudoh, J., Shimizu, N. & Narisawa, K. (1998). Molecular analysis of new Japanese patients with holocarboxylase synthetase deficiency. *J. Inher. Metabol. Dis.* **21**(8), 873-874.

Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C. & Wood, H. G. (1988). Evolutionary conservation among biotin enzymes. *J. Biol. Chem.* **263**, 6461-6464.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Schagger, H. & von Jagow, G. (1987). Tricine sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-79.

Schatz, P. J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Biotechnology* **11**(10), 1138-43.

Scrutton, M. C. & Taylor, B. L. (1974). Isolation and characterisation of pyruvate carboxylase from *Azotobacter vinlandii*. *Arch. Biochem. Biophys.* **164**, 641-654.

Sharma, R. C. & Schimke, R. T. (1996). Preparation of electrocompetent *E. coli* using salt free growth medium. *BioTechniques* **20**(1), 42-44.

Shenoy, B. C. & Wood, H. G. (1988). Effect of mutations at Met-88 and Met-90 on the biotination of Lys-89 of the apo 1.3S subunit of transcarboxylase. *FASEB J.* **2**, 2396-2401.

Shenoy, B. C., Xie, Y., Park, V. L., Kumar, G. K., Beegan, H., Wood, H. G. & Samols, D. (1992). The importance of methionine residues for the catalysis of the biotin enzyme, transcarboxylase. *J. Biol. Chem.* **267**(26), 18407-18412.

Sibler, A. P., Kempf, E., Glacet, A., Orfanoudakis, G., Bourel, D. & Weiss, E. (1999). *In vivo* biotinylated recombinant antibodies: high efficiency of labelling and application to the cloning of active anti-human IgG1 Fab fragments. *J. Immunol. Methods* **224**, 129-40.

Siegel, L., Foote, J. L. & Coon, M. J. (1965). The enzymatic synthesis of propionyl-coenzyme A holocarboxlyase from *d*-biotinyl 5'-adenylate and the apocarboxylase. *J. Biol. Chem.* **240**, 1025-1031.

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317.

Smith, J. L. (1998). Glutamine PRPP amidotransferase: Snapshots of an enzyme in action. *Curr. Opin. Struct. Biol.* **8**, 686-694.

Soumillion, P., Jespers, L., Bouchet, M., Marchand-Brynaert, J. & Fastrez, J. (1994). Phage display of enzymes and *in vitro* selection for catalytic activity. *Appl. Biochem. Biotechnol.* **47**(175-189).

Steinitz, F. (1898). *Arch. Gesamte Physiol.* **72**, 75.

Stolz, J., Darnhofer-Demar, B. & Sauer, N. (1995). Rapid purification of a functionally active plant sucrose carrier from transgenic yeast using a bacterial biotin acceptor domain. *FEBS Letts.* **377**, 167-171.

Stolz, J., Hoja, U., Meier, S., Sauer, N. & Schweizer, E. (1999). Identification of the plasma membrane H+-biotin symporter of *Saccharomyces cerevisiae* by rescue of a fatty acid-auxotrophic mutant. *J. Biol. Chem.* **274**(26), 18741-18746.

Stolz, J., Ludwig, A. & Sauer, N. (1998). Bacteriophage lambda surface display of a bacterial biotin acceptor domain reveals the minimal peptide size required for biotinylation. *FEBS Letts.* **440**, 213-217.

Streaker, E. D. & Beckett, D. (1998a). Coupling of site-specific DNA binding to protein dimerization in assembly of the biotin repressor biotin operator complex. *Biochemistry* **37**(9), 3210-3219.

Streaker, E. D. & Beckett, D. (1998b). A map of the biotin repressor-biotin operator interface - binding of a winged helix-turn-helix protein dimer to a forty base-pair site. *J. Mol. Biol.* **278**(4), 787-800.

Stucka, R., Dequin, S., Salmon, J. & Gancedo, C. (1991). DNA sequences in chromosomes II and VII code for pyruvate carboxylase isozymes in *Saccharomyces cerevisiae*: Analysis of pyruvate carboxylase deficient strains. *Mol. Gen. Genet.* **229**, 307-315.

Sumrada, R. A., Chisholm, G. & Cooper, T. G. (1982). Post-translational processing of urea amidolyase in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **2**, 800-804.

Sundaram, T. K., Cazzulo, J. J. & Kornberg, H. L. (1971). Pyruvate holocarboxylase formation from the apoenzyme and d-biotin in *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.* **143**, 609-616.

Sutton, M. R., Fall, R. R., Nervi, A. M., Alberts, A. W., Vagelos, P. R. & Bradshaw, R. A. (1977). Amino acid sequence of *Escherichia coli* biotin carboxyl carrier protein. *J. Biol. Chem.* **252**, 3934-3940.

Suzuki, Y., Aoki, Y., Ishida, Y., Chiba, Y., Iwamatsu, A., Kishino, T., Niikawa, N., Matsubara, Y. & Narisawa, K. (1994). Isolation and characterization of mutations in the human holocarboxylase synthetase cDNA. *Nat. Genet.* **8**(2), 122-128.

Takai, T., Yokayama, C., Wada, K. & Tanabe, T. (1988). Primary structure of chicken liver acetyl CoA carboxylase deduced from cDNA sequence. *J. Biol. Chem.* **263**, 2651-2657.

Taylor, B. L., Barden, R. E. & Utter, M. E. (1972). Identification of the reacting form of pyruvate carboxylase. *J. Biol. Chem.* **247**, 7383-7390.

Taylor, B. L., Frey, W. H., Barden, R. E., Scutton, M. C. & Utter, M. E. (1978). The use of the ultracentrifuge to determine the catalytically competent forms of enzymes with more than one oligomeric structure. Multiple reacting forms of pyruvate carboxylase from chicken and rat liver. *J. Biol. Chem.* **253**, 3062-3069.

Tissot, G., Douce, R. & Alban, C. (1997). Evidence for multiple forms of biotin holocarboxylase synthetase in Pea (*Pisum sativum*) and in *Arabidopsis thaliana* ; Subcellular fractionation studies and isolation of a cDNA clone. *Biochem. J.* **323**(1), 179-188.

Tissot, G., Job, D., Douce, R. & Alban, C. (1996). Protein biotinylation in higher plants; Characterization of biotin holocarboxylase synthetase activity from pea (*Pisum sativum*) leaves. *Biochem. J.* **314**(2), 391-395.

Tissot, G., Pepin, R., Job, D., Douce, R. & Alban, C. (1998). Purification and properties of the chloroplastic form of biotin holocarboxylase synthetase from *Arabidopsis thaliana* overexpressed in *Escherichia coli*. *Eur. J. Biochem.* **258**(2), 586-596.

Traub, W. (1956). *Nature* **178**, 649.

Val, D. L. (1995). The yeast pyruvate carboxylase 2 gene (Pyc2) and structure-function studies on yeast Pyc isozymes, University of Adelaide.

Val, D. L., Chapman-Smith, A., Walker, M. E., Cronan, J. E. & Wallace, J. C. (1995). Polymorphism of the yeast pyruvate carboxylase 2 gene and protein; Effects on protein biotinylation. *Biochem. J.* **312**(3), 817-825.

Varghese, J. N., Smith, P. W., Sollis, S. L., Blick, T. J., Sahasrabudhe, A., McKimmbreschkin, J. L. & Colman, P. M. (1998). Drug design against a shifting target - A structural basis for resistance to inhibitors in a variant of I nfluenza virus neuraminidase. *Structure* **6**(6), 735-746.

Vaughan, T. J., Osbourn, J. K. & Tempest, P. R. (1998). Human antibodies by design. *Nature Biotechnol.* **16**(6), 535-539.

Visser, C. M. & Kellogg, R. M. (1978). Biotin. Its place in evolution. *Mol. Evol.* **11**, 171-187.

Vrana, K. E. (1999). How the regulatory and catalytic domains get together. *Nat. Struct. Biol.* **6**, 442-448.

Walker, M. E., Jitrapakdee, S., Val, D. V. & Wallace, J. C. (1995). GenBank Accession number U30891. .

Walker, M. E., Val, D. L., Rhode, M., Devenish, R. J. & Wallace, J. C. (1991). Yeast pyruvate carboxylase; Identification of two genes encoding isoenzymes. *Biochem. Biophys. Res. Comm.* **176**, 1210-1217.

Wallace, J. C. & Eastervrook-Smith, S. B. (1985). The structure of pyruvate carboxylase. In *Pyruvate carboxylase* (Keech, D. B. & Wallace, J. C., eds.), pp. 66-108. CRC Press, Boca Raton, FL.

Wallace, J. C., Jitrapakdee, S. & Chapman-Smith, A. (1998). Pyruvate carboxylase. *Int. J. Biochem. Cell Biol.* **30**(1), 1-5.

Wallace, P. G., Pedler, S. M., Wallace, J. C. & Berry, M. N. (1994). A method for the determination of the cellular phosphorylation potential and glycolytic intermediates in yeast. *Anal. Biochem.* **222**(2), 404-408.

Wallis, N. G., Allen, M. D., Broadhurst, R. W., Lessard, I. A. D. & Perham, R. N. (1996). Recognition of a surface loop of the lipoyl domain underlies substrate channelling in the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **263**, 463-474.

Wallis, N. G. & Perham, R. N. (1994). Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**(209-216).

Wang, Z. X., Bhargava, A., Sarkar, R. & Gerino, F. J. (1996). Vectors for a double tagging assay for protein-protein interactions: localisation of the CDK2-binding domain of human p21. *Gene* **173**, 147-154.

Whitney, P. A. & Cooper, T. G. (1972). Urea carboxylase and allophanate hydrolase: Two components of ATP:urea amido-lyase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **247**, 1349-1353.

Wilchek, M. & Bayer, E. A. (1990). Avidin - biotin technology. *Methods Enzymol.* **184**.

Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. (1992). *Escherichia coli* biotin holoenzyme synthetase/*bio* repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc. Natl. Acad. Sci. USA* **89**, 9257-9261.

Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. (1994). Making antibodies by phage display. *Ann. Rev. Immunol.* **12**, 433-455.

Wolf, B. (1995). Disorders of biotin metabolism. In *The metabolic and molecular basis of inherited diseases* (Scriver, C. R., Beavder, A. L. & Sly, W. S., eds.), Vol. 7, pp. 3151-3177. McGraw-Hill, New york.

Wolf, D. E., Valiant, J., Peck, R. L. & Folkers, K. (1952). *J. Am. Chem. Soc.* **74**, 2002.

Wood, H. G. & Barden, R. E. (1977). Biotin enzymes. *Ann. Rev. Biochem.* **46**, 385-413.

Wood, H. G., Harmon, F. R., Wuhr, B., Hubner, K. & Lynen, F. (1980). Comparison of the biotination of apotranscarboxylase and its aposubunits. Is assembly essential for biotination? *J. Biol. Chem* **255**, 7397-7409.

Wood, V., Rajandream, M. A., Barrell, B. G., Lauber, J., Hilbert, H. & Duesterhoeft, A. (1997). Accession Number O14353. *EMBL/GenBank/DDBJ databases.*

Xia, W.-L., Zhang, J. & Ahmad, F. (1994). Biotin holocarboxylase synthetase: Purification from rat liver cytosol and some properties *Biochem. Mol. Biol. Int.* **34**(225-232).

Xu, X., Matsuno-Yagi, A. & Yagi, T. (1993). DNA sequencing of the seven remaining structural genes of the gene cluster encoding the energy transducing NADH-Quinone oxidoreductase of *Paracoccus denitrificans.* *Biochemistry* **32**, 968-981.

Xu, Y. & Beckett, D. (1994). Kinetics of biotinyl-5'-adenylate synthesis catalyzed by the *Escherichia coli* repressor of biotin biosynthesis and the stability of the enzyme-product complex. *Biochemistry* **33**(23), 7354-7360.

Xu, Y. & Beckett, D. (1996). Evidence for interdomain interaction in the *Escherichia coli* repressor of biotin biosynthesis from studies of an N-terminal domain deletion mutant. *Biochemistry* **35**(6), 1783-1792.

Xu, Y. & Beckett, D. (1997). Biotinyl-5'-adenylate synthesis catalysed by *Escherichia coli* repressor of biotin synthesis. *Methods Enzymol.* **279**, 405-421.

Xu, Y., Johnson, C. R. & Beckett, D. (1996). Thermodynamic analysis of small ligand binding to the *Escherichia coli* repressor of biotin biosynthesis. *Biochemistry* **35**(17), 5509-5517.

Xu, Y., Nenortas, E. & Beckett, D. (1995). Evidence for distinct ligand-bound conformational states of the multifunctional *Escherichia coli* repressor of biotin biosynthesis. *Biochemistry* **34**(51), 16624-16631.

Yao, X., Soden, C., Summers, M. F. & Beckett, D. (1999). Comparison of the backbone dynamics of the apo- and holo-carboxy-terminal domain of the biotin carboxyl carrier subunit of *Escherichia coli* acetyl-CoA carboxylase. *Prot. Sci.* **8**(2), 307-317.

Yao, X., Wei, D., Soden, C., Summers, M. F. & Beckett, D. (1997). Structure of the carboxy-terminal fragment of the apo-biotin carboxyl carrier subunit of *Escherichia coli* acetyl-CoA Carboxylase. *Biochemistry* **36**(49), 15089-15100.

Zempleni, J., Trusty, T. A. & Mock, D. M. (1997). Lipoic acid reduces the activities of biotin-dependent carboxylases in rat liver. *J. Nutr.* **127**(9), 1776-1781.

Zhang, S., Sanyal, I., Bulboaca, G. H., Rich, A. & Flint, D. H. (1994). The gene for biotin synthase from *Saccharomyces cerevisiae*: Cloning, sequencing and complementation of *Escherichia coli* strains lacking biotin synthase. *Arch. Biochem. Biophys.* **309**, 29-35.