Functional characterisation of the N-terminal region of Holocarboxylase synthetase

by

Lungisa Mayende, M.Sc.

A thesis submitted to the University of Adelaide, South Australia in fulfilment of the requirements for the degree of Doctor of Philosophy

May, 2012

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

Abstract

Biotin (vitamin H or B7) is an important micronutrient that is covalently attached to biotin-dependent enzymes by human biotin protein ligase (hBPL) or holocarboxylase synthetase (HCS). Patients with HCS deficiency are treated with oral biotin supplementation, which in most cases is able to reverse the clinical symptoms. However, some patients respond poorly to biotin therapy and have an extremely poor long-term prognosis. The molecular explanation for this is not understood. In this study HCS was investigated to improve our understanding of this key enzyme.

The catalytic region of all BPLs is contained in the conserved C-terminal region. HCS contains a long N-terminal extension that is not present in bacterial BPLs. The structure and function of the N-terminal region is yet to be determined. In order to delineate the domain structure of HCS limited proteolysis was performed previously in our laboratory. Two protease-sensitive linker regions were identified, one between residues 151-153, the other at amino acid 314. Of particular importance is the proposed structured domain containing residues 159-314, as amino acid substitutions in this region have been shown to compromise enzyme activity despite being distal to the C-terminally located active site. This thesis provides genetic evidence for a direct interaction between the N-terminal and C-terminal halves of HCS using a yeast two-hybrid assay. This interaction was mapped using a trucation study to the proposed structured domain of the N-terminal region of HCS (159-314 N-HCS).

HCS deficiency gives rise to the metabolic disorder multiple carboxylase deficiency (MCD). Mutations within the proposed structured domain 159-314 N-HCS give rise to MCD patients that are poorly responsive to the current therapy (biotin supplementation) and have an extremely poor long-term prognosis. In this thesis, a series of novel mutations in the

proposed structured domain 159-314 N-HCS were generated using “error prone” PCR. The catalytically inactive mutants were isolated from the library using an *in vivo* complementation assay. The mutants that were isolated were identified by DNA sequencing as L166R, L206P, W210R, L246M, L270S, H306R, F321S and the double mutant E181G, E327G. These residues are highly conserved within vertebrate species. These novel HCS mutants, together with the MCD mutants L216R HCS and L237P HCS, were employed to further characterise the function of the proposed structured N-terminal domain. Using the Yeast Two-hybrid assay, it was shown that the interaction between the two halves of HCS was not disrupted by the MCD mutants nor the novel mutants. Conversely, it was shown that the MCD mutants, and the majority of the novel HCS mutants, disrupted the interaction between HCS and its protein substrate the Pyruvate Carboxylase biotin domain hPC107.

Surface Plasmon Resonance was then employed to further characterise this observation. This study has demonstrated for the first time that although the association between HCS and its substrate was not compromised by mutation, the MCD mutants had a >15-fold increase in dissociation rate from the substrate compared to wild type HCS. This work provided a novel function for the proposed structural N-terminal domain. Furthermore, these data provide a molecular explanation for the HCS deficient patients that do not respond to biotin therapy.

Statement of Originality

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

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

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

Signed

…….....

Lungisa Mayende

Acknowledgements

I would like to thank my supervisor, Grant Booker for having a place for me to work in his lab. Before I came to Australia your enthusiasm was encouraging and I am glad that I completed my PhD in your lab. Thank you for your expertise, knowledge and perfectionism that have helped to produce this thesis.

I extend gratitude to my co-supervisor, Steven Polyak. I appreciate the amount of time and support that you gave me these past few years. This thesis would not have been completed without your constant and untiring advice, guidance, expertise and assistance.

I was the recipient of the International Postgraduate Research Scholarship (IPRS) and the Adelaide Scholarships International (ASI), which funded my PhD studies. I appreciated the financial assistance.

I would like to thank the Booker Lab members for their support during my PhD. Especially Tatiana Soares da Costa, you have been a great friend (more like a sister, you are so right), someone to bounce ideas off of and helpful (thanks for assisting me with SPR). I am grateful to have such a beautiful person in my life, thank you for choosing this lab! Thank you Kate Wegener for helping me with NMR analysis and general structural work. Thank you Colleen Bindloss for your assistance with gel filtration and ASTRA. Thank you to James Eddes and Chris Crusaro from the Adelaide Proteomics Centre for your assistance with the MS work. Thanks to Al Azhar and Wanisa Salaemae for making the lab the positive place that it is. I extend thanks to previous Booker Lab members especially Ethan Chen (we made it, you're next!), Iain Murchland (thanks for the helpful structural models discussion), Rachel Swift (thanks for the work on HCS that you performed) and Shee Chee Ong (thanks for the cloning you did). I would also like to thank previous BPL team

members, namely Lisa Bailey (your work on HCS was invaluable) and Nicole Pendini. Thanks to the Wallace lab for being our go to place when we need to borrow something (we always return the favour).

To everyone in the department, the different labs and the lovely people from CSU and other support staff, thank you for the smiles and chats in the hallways. You make the MBS building what it is.

I would like to thank all of my friends and family that have supported me these past few years. I love you each and every one of you! Thank you Mama for your love and belief in me. Being away from you has been hard but you have supported all of my dreams, including this one. Your constant encouragement gave me strength. I love you, my twin soul. To my partner Shane Devitt, thanks for being my best friend. For being the person I turned to when I was frustrated, needed support or needed a hug. You always found a way to make me smile, you uplifted me. I stayed sane these past few years because of you. Thank you for also sharing in my triumphs. I love you boo! :)

And finally, thank you God for getting me through to the other side of this thesis. Without you pulling me together, I would have surely fallen apart. I am very content to have made it!

Table of Contents

Abstract.....**ii**
Statement of Originality.....**iv**
Acknowledgements.....**v**
Table of Contents.....**vii**
List of Publications.....**xi**
List of Abbreviations.....**xiii**

Chapter 1 Introduction.....**1**
1 Introduction.....2
1.1 Biotin and biotin-dependent enzymes.....2
1.2 Biotin domains.....4
1.3 Biotin protein ligase.....7
1.4 Holocarboxylase synthetase (HCS).....8
1.5 Microbial BPLs provide structural and functional information on HCS.....15
1.5.1 Class I BPLs.....15
1.5.1.1 *Pyrococcus horikoshii* OT3 BPL (*Ph*BPL).....15
1.5.2 Class II BPLs.....21
1.5.2.1 BirA.....21
1.5.3 Class III BPLs.....23
1.5.3.1 Yeast biotin protein ligase (yBPL).....23
1.6 Comparison of the N-terminal and C-terminal regions of HCS (and other Eukaryotic BPLs) with N- and C- terminal regions of microbial BPLs.....24
1.7 Multiple Carboxylase Deficiency (MCD).....27
1.7.1 K_M MCD mutants.....30
1.7.2 V_{MAX} MCD mutants.....31

Chapter 2 General Materials and Methods.....**36**
2.1 Materials.....37
2.1.1 General materials.....37
2.1.2 Chemical reagents.....37
2.1.3 Restriction endonucleases.....38
2.1.4 Antibodies.....38
2.1.5 Bacterial strains.....38
2.1.6 Bacterial media.....38
2.1.7 Yeast Strain.....39
2.1.8 Yeast media.....39
2.1.8.1 Synthetic-Dropout Amino Acid Mix.....40

2.1.9 Oligonucleotide Primers…..…….....41
2.1.10 Cloning vectors……..…..42
2.1.10.1 Bacterial Cloning Vectors and Protein Expression Vectors....42
2.1.10.2 Yeast Two-Hybrid Vectors……42
2.1.11 Commercial kits……..…..43
2.1.12 Buffers and solutions……..43
2.1.13 Plasmids and vectors……44
2.1.14 Computer software……..…44
2.1.15 Web resources……..…44
2.2 General methods……45
2.2.1 Determination of protein concentration……..…45
2.2.2 Purification of GST-tagged proteins using Profinia™……....45
2.2.3 SDS-PAGE electrophoresis and gel staining……..…45
2.2.4 Western Blotting……..…46
2.2.5 HCS activity assays……..…46
2.2.6 Circular Dichroism……47
2.2.7 Molecular biology techniques……..47
2.2.7.1 Primer Design……..…47
2.2.7.2 PCR protocols……..…47
2.2.7.3 Site-Directed Mutagenesis……..…..48
2.2.7.4 Phosphorylation and annealing of Oligonucleotides…….....…48
2.2.7.5 Agarose gel electrophoresis……..…49
2.2.7.6 Preparation of DH5α competent cells……..……..49
2.2.7.7 Restriction digest of DNA……..50
2.2.7.8 Ligation of DNA fragments……....50
2.2.7.9 Transformation of competent cells……..…….50
2.2.7.10 Glycerol stocks……..……50
2.2.7.11 Purification of plasmid DNA……..……50
2.2.7.12 Quantification of DNA……..…51
2.2.7.13 DNA sequencing……..…51
2.2.8 Yeast methods……..……52
2.2.8.1 Preparation of competent *S. cereviseae*……..……52
2.2.8.2 Transformation of competent yeast strains……..…52

Chapter 3 Exploring the inter-domain interactions within HCS……..**53**
3.1 Introduction……54
3.1.1 Investigating the inter-domain interactions within HCS ……..…54
3.2 Specific Methods……..…56
3.2.1 Nucleic acid manipulations……....56
3.2.2 *In vivo* selection of HCS mutants……...56

3.2.3 Yeast Methods……..…57
3.3 Results……...58
3.3.1 Determination of an interaction between the N-terminal and C-terminal regions of HCS using the Yeast Two-hybrid assay……58
3.3.2 Constructs prepared for the interaction studies using Yeast Two-hybrid assay……..…59
3.3.3 Defining the region of N-HCS that interacts with C-HCS……..61
3.3.4 *In vivo* complementation assay……..…63
3.3.5 Determining the effect of MCD and novel HCS mutants on the interaction between N-HCS and C-HCS domains……..…66
3.3.6 Determining the effect of MCD and novel HCS mutants on the intermolecular interaction of HCS and hPC107……..….69
3.4 Discussion……..…72

Chapter 4 NMR studies of the N-terminal cap of HCS ……...…**75**
4.1 Introduction……76
4.1.1 Structural Determination of proteins by NMR Spectroscopy…..……76
4.2 Specific Methods……..79
4.2.1 Cloning of pGEX NTC, the NTC variants pGEX NTC (G159-E293) and pGEX NTC (G159-D349)……...79
4.2.2 Transformation, expression and purification of GST-NTC (G159-R314), GST-NTC_2 (G159-E293) and GST-NTC_3 (G159-D349)……80
4.2.3 Cloning, transformation and expression of NTC using alternative expression vectors……..……...…80
4.2.4 Purification of NTC using on column cleavage……..……....80
4.2.5 2D NMR Spectroscopy Experiments……81
4.2.6 Investigation of the secondary structure composition, structural integrity and thermal stability of NTC using Circular dichroism…….…….82
4.2.7 Investigation of the oligomeric state of NTC using gel filtration chromatography……..83
4.2.8 Molecular structural models of 159-314 N-HCS or NTC and 1-151 N-HCS……....83
4.3 Results and Discussion……...84
4.3.1 Selection of NTC construct……84
4.3.2 Expression and purification of NTC (sample 1)……..……85
4.3.3 Expression and purification of NTC (sample 2)……..……86
4.3.4 Expression and purification of NTC (sample 3), indentification of NTC using Mass Spectrometry (MS) and mass determination of NTC using MALDI MS…….…..87

4.3.5 2D NMR experiments of NTC using a triple resonance $^{1}H/^{13}C/^{15}N$ PFG probe (sample 1) and triple resonance $^{1}H/^{13}C/^{15}N$ PFG cold probe (sample 2)……..…90
4.3.6 Circular dichroism……..…94
4.3.6.1 CD analysis of NTC…….94
4.3.6.2 Analysis of NTC CD spectrum using CONTINLL, CDSSTR and SELCON3……..……96
4.3.7 Investigation of the oligomeric state of NTC using gel filtration chromatography……98
4.3.8 *De novo* predictions for the structures of NTC or 159-314 N-HCS and 1-151 N-HCS……100
4.4 Summary and Future work……..…106

Chapter 5 The Characterisation of MCD HCS mutants…….……**110**
5.1 Introduction……..…111
5.2 Specific Methods……....112
5.2.1 Expression and purification of hPC107……..112
5.2.2 Expression and purification of wt HCS, L216R HCS and L237P HCS..112
5.2.3 Circular Dichroism (CD) Spectroscopy……...113
5.2.4 Surface Plasmon Resonance (SPR)……..114
5.3 Results…….115
5.3.1 Expression and purification of hPC107……..117
5.3.2 Expression and purification of wt HCS, L216R HCS and L237P HCS..118
5.3.2.1 Optimisation of HCS purification using a higher concentration of imidazole……...118
5.3.3 HCS activity assay……....120
5.3.4 Circular Dichroism (CD) Spectroscopy……...…120
5.3.5 Surface Plasmon Resonance (SPR)……...……..122
5.3.5.1 pH scouting and immobilisation of protein ……122
5.3.5.2 Determination of the activity of immobilised HCS on CM5 chip and binding of biotin and MgATP occurs in an ordered manner…..…123
5.3.5.3 Binding of biotinol-5´-AMP to wild type and mutant HCS…...127
5.3.5.4 Binding hPC107 to wild type and mutant HCS……...…128
5.4 Discussion……..……..…133
6 Final Discussion……..…**136**
6.1 Functional characterisation of the N-terminal region of HCS……..……137
6.2 A novel molecular mechanism to explain the juvenile metabolic syndrome biotin-unresponsive HCS deficiency……..…138
6.3 New directions for the development of V_{MAX} HCS deficiency therapeutics……...139
References……..……..…**144**

List of Publications

Pubished Manuscript

Lungisa Mayende, Rachel D. Swift, Lisa M. Bailey, Tatiana P. Soares da Costa, John C. Wallace, Grant W. Booker and Steven W. Polyak. (2012) A novel molecular mechanism to explain biotin-unresponsive holocarboxylase synthetase deficiency, *J. Mol Medicine*, 90(1): 81-88.

Communications and Conference Proceedings

Mayende, L., Swift, R.D., Bailey, L.M. Wallace, J.C., Booker, G.W. & Polyak, S.W. (2010) Characterisation of the N-terminal region of Holocarboxylase synthetase. Ozbio Conference, Poster 255.

Mayende, L., Swift, R.D., Bailey, L.M., Wallace, J.C., Booker, G.W. & Polyak, S.W. (2010) Mechanism to explain the juvenile metabolic syndrome Multiple Carboxylase Deficiency. Australian Society for Medical Research SA division Scientific Meeting, Poster 58.

Mayende, L., Swift, R.D., Bailey, L.M. Wallace, J.C., Booker, G.W. & Polyak, S.W. (2010) Domain mapping and functional analysis of the N-terminal region of Holocarboxylase synthetase. Adelaide Protein Group (APG) Student Award **Finalist**, Oral Presentation.

Mayende, L., Swift, R.D., Bailey, L.M. Wallace, J.C., Booker, G.W. & Polyak, S.W. (2010) Domain mapping and functional analysis of the N-terminal region of Holocarboxylase synthetase. 35th Lorne Conference on Protein Structure and Function, Poster 285. **Awarded** student travel award.

Mayende, L., Swift, R. D., Bailey, L. M., Pendini, N. R., Wallace, J. C., Booker, G.W. & Polyak, S. W. (2010) Protein biotinylation: An example of a post-translational modification of exceptional specificity. 3rd Annual Protein and Peptide Conference (Pep Con). Beijing, China.

Mayende, L., Swift, R.D., Bailey, L.M., Pendini, N.R., Wallace, J.C., Booker, G.W. & Polyak, S.W. (2008) Interaction and Mutagenesis studies on Holocarboxylase synthetase (HCS). Australian Society for Medical Research SA division Scientific Meeting, Oral Presentation.

Mayende, L., Swift, R. D., Bailey, L. M., Pendini, N. R., Wallace, J. C., Polyak, S. W. & Booker, G.W. (2008) Domain structure of human Holocarboxylase synthetase: evidence of an interaction between the N-terminal and C-terminal halves. 33rd Lorne Conference on Protein Structure and Function, Poster 255. **Awarded** student travel award.

Mayende, L., Swift, R. D., Bailey, L. M., Pendini, N. R., Wallace, J. C., Polyak, S. W. & Booker, G.W. (2007) Evidence of an interaction between the N-terminal and C-terminal halves of Holocarboxylase synthetase. University of Adelaide, School of Molecular and Biomedical Sciences Research Symposium.

List of Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
Ab	antibody
ACC	acetyl CoA carboxylase
AD	activation domain
Amp	ampicillin
AMP	adenosine monophosphate
AP	alkaline phosphatase
Apo	unliganded enzyme
ATP	adenosine triphosphate
BirA	biotin inducible repressor A
BCCP	biotin carboxyl carrier protein
BLAST	basic local alignment search tool
BME	beta-mercaptoethanol
bp	base pair
BPL	biotin protein ligase
BSA	bovine serum albumin
C-	carboxyl-
cDNA	complementary deoxyribonucleic acid
C-HCS	C-terminal amino acids 315-726 of Holocarboxylase synthetase
D_2O	deuterium oxide
DMSO	dimethyl sulfoxide
DNA	deoxynucleotide triphosphate
DNA BD	DNA binding domain
dNTPs	deoxynucleotide triphosphates
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetra-acetic acid
GST	glutathione-S-transferase
HCS	Holocarboxylase synthetase
Holo	ligand bound enzyme
hPC107	107 amino acids encoding the biotin domain of human pyruvate carboxylase

hr	hour
HRP	horseradish peroxidase
IP	immunoprecipitation
k_a	association rate constant
k_d	dissociation rate constant
K_D	equilibrium dissociation constant
K_M	Michaelis constant
kb	kilobase pair
kDa	kilodalton
LB	Luria broth
LiAc	Lithium Acetate
m	metre
M	molar
µ	micron
mA	milliampere
min	minute, minutes
MCC	methylcrotonyl-CoA carboxylase
MCD	multiple carboxylase deficiency
MOPS	3-morpholinopropanesulfonic acid
*Mt*BPL	*Mycobacterium tuberculosis* BPL
MW	molecular weight
MWCO	molecular weight cut-off
n	nano
N-	amino-
N-HCS	N-terminal amino acids 1-314 of Holocarboxylase synthetase
NMR	nuclear magnetic resonance
ODx	nm optical density at x nm wavelength
PCC	propionyl-CoA carboxylase
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline and 0.05% (v/v) Tween-20
PC	pyruvate carboxylase
PCR	polymerase chain reaction
PDB	protein data bank
*Ph*BPL	*P.horikoshii* BPL
PMSF	phenylmethylsulfonylfluoride

PVDF	polyvinyl difluoride
RMSD	root mean square deviation
RNA	ribonucleic acid
RT	room temperature
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
s	second
TBS	tris buffered saline
TBS-T	tris buffered saline and 0.1% (v/v) Triton-X
Tris	2-amino-2-hydroxymethylpropane-1,3-diol
Tween-20	polyoxyethylene-sorbitan monolaurate
U	units
V_{MAX}	maximum velocity
UV	ultra violet
WT	wild type
yBPL	Yeast (*S.cerevisiae*) biotin protein ligase

CHAPTER 1

Introduction

1 Introduction

1.1 Biotin and biotin-dependent enzymes

Biotin, also known as vitamin H and vitamin B-7, which was discovered by Kogl and Tonnis in 1934, is a water soluble vitamin found in all organisms. It is synthesised by plants, most prokaryotes and some fungi. Higher organisms, such as mammals, are incapable of synthesising biotin and therefore acquire the vitamin from their diet (Pacheco-Alvarez *et al.*, 2002). Dietary biotin is in either the free or protein bound forms. The protein bound form of biotin, biocytin (biotinyl-ε-lysine), is hydrolysed by the enzyme biotidinase and biotin is thereby recycled (Hymes and Wolf, 1999). The free biotin acts as a cofactor of biotin-dependent carboxylases, decarboxylases and transcarboxylases. Biotin is physiologically active when it is covalently attached to the ε-amino group of a specific lysine residue at the active site of these enzymes (Chapman-Smith *et al.*, 1999; Campeau and Gravel, 2001). The enzyme responsible for biotin attachment is biotin protein ligase (BPL). In mammals this enzyme is commonly known as Holocarboxylase synthetase and is the subject of this thesis.

Biotin-dependent enzymes catalyse important metabolic processes including fatty acid synthesis, gluconeogenesis, amino acid catabolism and energy transduction. These enzymes employ the biotin prosthetic group as a mobile carboxyl carrier and catalyse the transfer of carbon dioxide to and between metabolites (Samols *et al.*, 1988; Knowles, 1989). In *Escherichia coli*, acetyl-CoA carboxylase is the only biotin-dependent enzyme whereas mammals have five biotin-dependent enzymes (figure 1.1). These enzymes include acetyl-CoA carboxylase (isoforms ACC 1 and ACC 2), methylcrotonyl-CoA carboxylase (MCC), propionyl-CoA carboxylase (PCC) and pyruvate carboxylase (PC) (Campeau and Gravel, 2001).

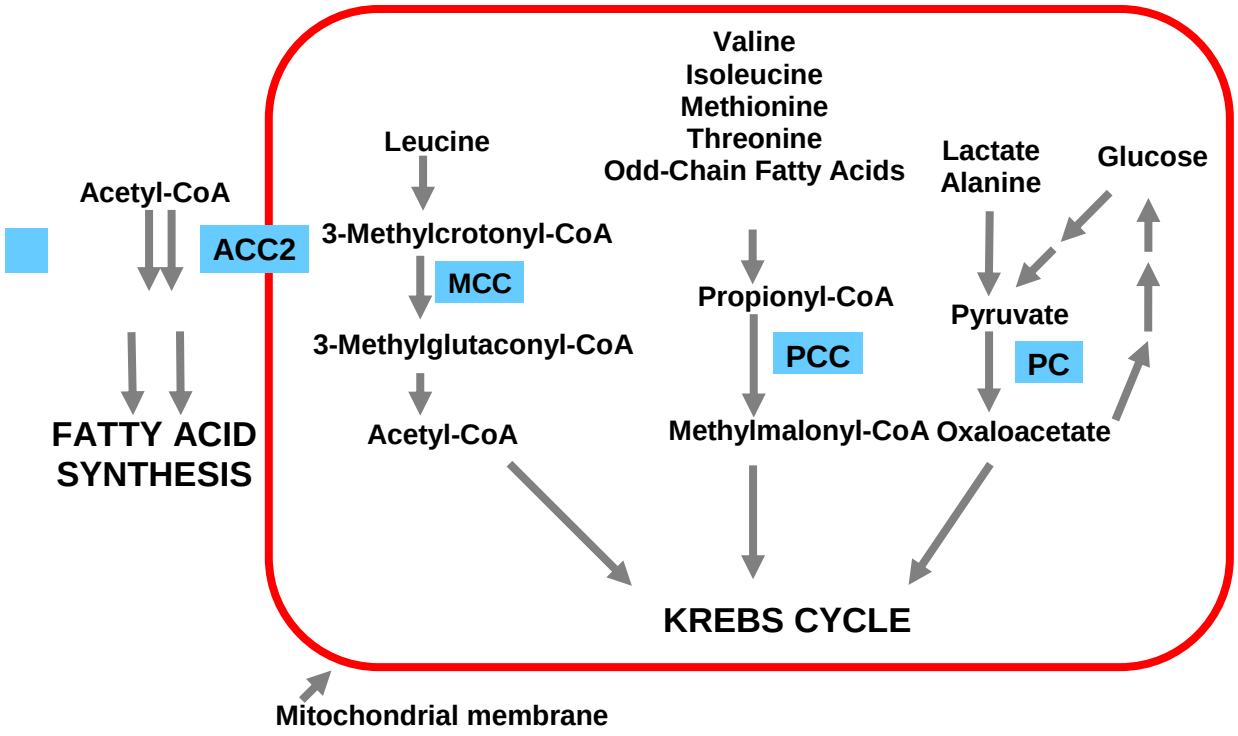


Figure 1.1 Role of biotin-dependent carboxylases in human metabolism. Adapted from Pacheco-Alvarez *et al.* (2002).

The mammalian biotin-dependent enzymes are essential as they are involved in fatty acid synthesis, gluconeogenesis and amino acid catabolism (Wolf, 1995). There are two isoforms of ACC that catalyse the carboxylation of acetyl CoA to form malonyl CoA and is required in the biosynthesis of fatty acids (Pacheco-Alvarez *et al.*, 2002). Whilst both ACC1 and ACC2 have enzymatic activities in the cytoplasm ACC1 is soluble whereas ACC2 is anchored to the outer membrane of the mitochondrion. It is believed that malonyl CoA, produced by ACC1, is used in fatty acid synthesis. ACC2 regulates the movement of fatty acids in the mitochondrion for β oxidation by producing malonyl CoA, which is an allosteric inhibitor of carnitine palmitoyltransferase-1 (CPT-1), the enzyme involved in transport of cytosolic long-chain acyl CoA molecules into the mitochondria. The other three biotin enzymes are located in the mitochondria. MCC catalyses the carboxylation of 3-methylcrotonyl-CoA to 3-methylglutaconyl-CoA, and thereby is involved in the catabolism

of leucine (McKean *et al.,* 2000). PCC catalyses the conversion of propionyl-CoA to methylmalonyl-CoA and therefore is involved in the catabolism of branched chain amino acids and fatty acids of odd-numbered chain length (Pacheco-Alvarez *et al.*, 2002). PC catalyses the reaction of pyruvate to oxaloacetate, which is an intermediate in the Krebs cycle. Additionally oxaloacetate is an intermediate in phosphoenolpyruvate synthesis and is therefore important for the synthesis of glucose (Wallace *et al.*, 1998).

1.2 Biotin domains

The attachment of biotin to biotin-dependent enzymes is catalysed protein ligase (BPL). Biotin attachment occurs at a specific lysine residue within a highly conserved biotin domain structure found in all biotin-dependent enzymes. Biotin carboxyl carrier protein (BCCP), a subunit of the *E.coli* acetyl CoA carboxylase, was the first characterized 3D structure of a biotin domain (Athappilly and Hendrickson, 1995). Both apo and holo forms of the BCCP were solved using X-ray crystallography and NMR spectroscopy (Athappilly and Hendrickson, 1995; Roberts *et al.*, 1999, figure 1.2). The protein:protein interaction between BPLs and biotin enzymes is conserved, as it has been shown that BPLs are interchangeable between organisms (Chapman-Smith and Cronan, 1999; Chapman-Smith *et al.*, 1999). This is possibly due to the structural conservation between the biotin domains (Pendini *et al.*, 2008) and the conserved reaction catalysed by BPL.

The biotin domain of BCCP consists of a barrel structure containing two antiparallel β-sheets, each containing four strands. The N- and C- termini are in close proximity at one end and the biotinyl-lysine exposed on a tight β-turn on the opposite face of the molecule (Reche *et al.*, 1998; Roberts *et al.*, 1999). The correct positioning of the lysine residue on the hairpin loop is critical, as moving it by just one position to the N- or C-terminal side abolishes biotinylation *in vitro* and *in vivo* (Reche *et al.*, 1998).

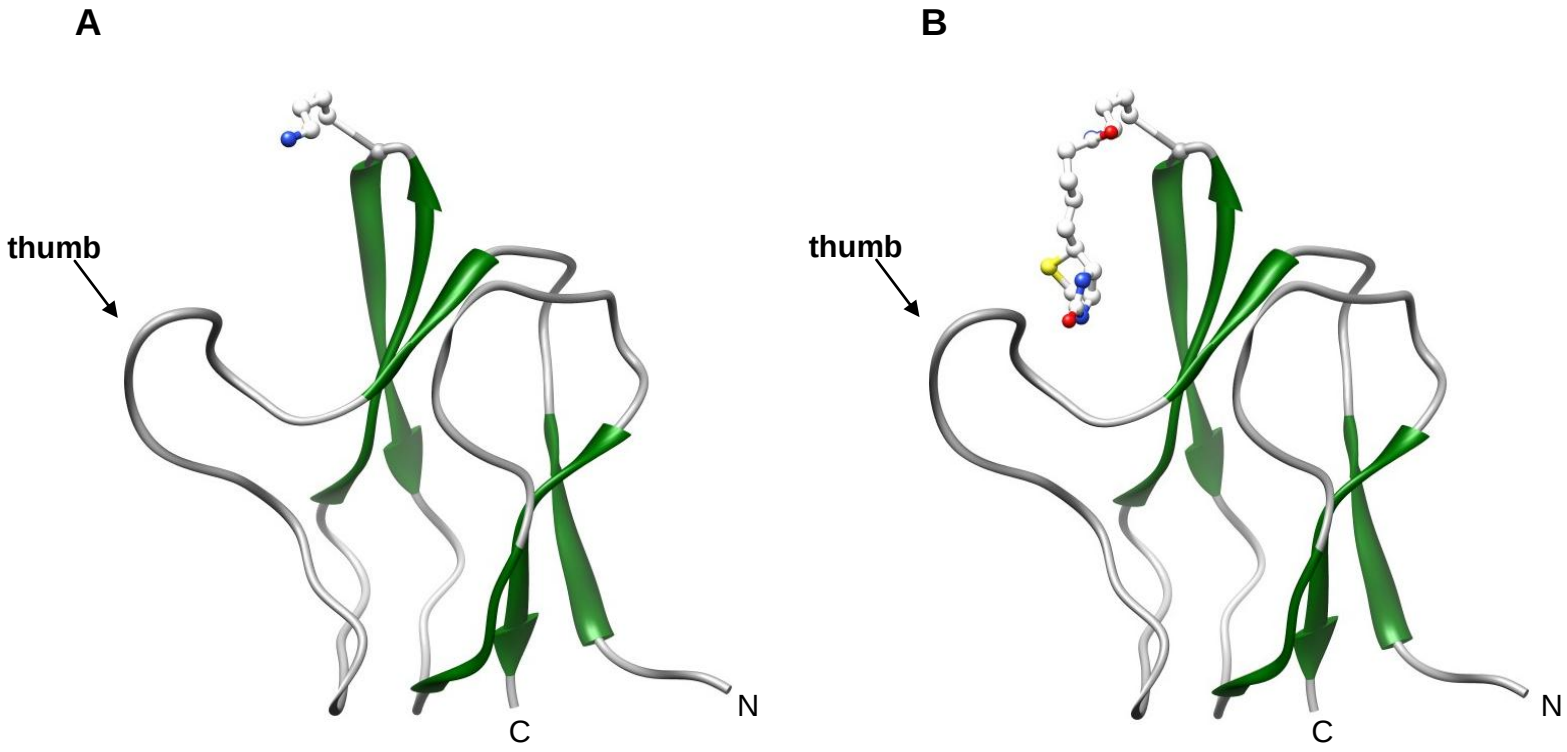


Figure 1.2 The X-ray crystal structure of the biotin carboxyl carrier protein of *Escherichia coli* showing the target lysine 122, viewed in ball and stick at the exposed hairpin loop of β strands 4 and 5. **A** is apo BCCP and **B** is holo BCCP (Athappilly and Hendrickson, 1995). Figure generated using UCSF Chimera using PDB 1DBO.

BPLs and lipoate protein ligases form the highly divergent enzyme protein family PFAM 03099 (Christensen and Cronan, 2010). Despite this divergence BCCP is structurally analogous to the lipoyl domain, where lipoic acid is attached to a specific lysine residue within a highly conserved DKA motif of lipoic acid dependent enzymes (Reche *et al.*, 1999). Mutation of the MKM biotinylation sequence in BCCP to the DKA motif resulted in loss of biotinylation and a low level of lipoylation *in vitro* (Reche *et al.*, 1998). A structural cue that prevents lipoylation of BCCP in *E.coli* is the presence of a "protruding thumb" structure (figure 1.2), which consists of 8 amino acids between the β2 and β3 strands (Reche *et al.*, 1999). Amino acid sequence alignments suggest that this feature is absent in eukaryotic biotin domains. The mechanism eukaryotic BPLs utilise to distinguish the target lysine in the biotin domain from the lipoyl domain is currently unknown. There is a high degree of sequence similarity and functional conservation of biotin domains between biotin dependent enzymes and between species. BPLs from bacteria can biotinylate biotin domains from eukaryotic substrates and vice versa (Chapman-Smith *et al.*, 1999; Samols *et al.*, 1988). A study by Healy *et al.* (2010) determined that human BPL or Holocarboxylase synthetase (HCS) has a reduced activity toward bacterial BCCP87 relative to its substrate p67 (which is the C-terminal 67 amino acid BCC domain of the PCC α subunit). They found that BirA had similar biotinylation activity toward the two substrates. When they engineered a "thumbless" bacterial BCCP, it could be biotinylated by HCS with similar activity to that of p67. The study suggested that the thumb loop found in bacterial carboxylases interferes with biotinylation by HCS.

1.3 Biotin protein ligase

Biotin protein ligase (BPL) catalyses the attachment of biotin to the lysine residue of biotin-dependent enzymes. This reaction is carried out in two steps, which are outlined in figure 1.3. The BPL reaction results in the formation of an amide linkage between the carboxyl group of biotin and the ε-amino group of the target lysine in the biotin accepting domain, in an ATP-dependent manner (Chapman-Smith *et al.*, 1999; Campeau and Gravel, 2001).

biotin + ATP ↔ biotinyl-5´-AMP +PPi (1)

biotinyl-5´-AMP + apo-protein → holo-protein + AMP (2)

Figure 1.3 The biotin protein ligase (BPL) reaction

The mechanism of the BPL reaction is analogous to that of the aminoacyl-tRNA synthetases whereby the reaction proceeds through an adenylated intermediate. In the first step biotin reacts with ATP, resulting in the formation of the biotinyl-5´-AMP intermediate. Biotinyl-5´-AMP stays bound to the active site until a suitable substrate is present (Chapman-Smith *et al.*, 1999). In the second step the nucleophilic ε-amino group of the lysine attacks the mixed anhydride carbon atom to form a covalent amide bond between the biotin and the lysine side chain with AMP as the product. Once the amide bond is formed, the biotin-dependent enzyme remains biotinylated for the remainder of its lifecycle (Chapman-Smith *et al.*, 1999).

BPLs fall into four structural classes (figure 1.4). Class I BPLs contain mono-functional BPLs that only consist of the catalytic core. *Pyrococcus horikoshii* BPL (*Ph*BPL) and *Mycobacterium tuberculosis* BPL (*Mt*BPL) are examples from this class. Class II represents a bi-functional BPL containing a N-terminal Helix-Turn-Helix (HTH) domain, which is

required for DNA binding and contributes to the transcriptional repressor function of BPL (Xu and Beckett, 1994). *E. coli* BPL (*Ec*BPL or BirA) and *Staphylococcus aureus* BPL (*S. aureus* BPL) are examples of Class II BPLs. An extended N-terminal region and absence of HTH domain in Class III is characteristic for eukaryotic BPLs. The N-terminal domain is required for catalysis by both yeast BPL (Polyak *et al.*, 1999) and HCS (Lee *et al.*, 2010). Plant BPLs contain an organelle targeting sequence in the N-terminal region (Tissot *et al.*, 1997) and are Class IV BPLs.

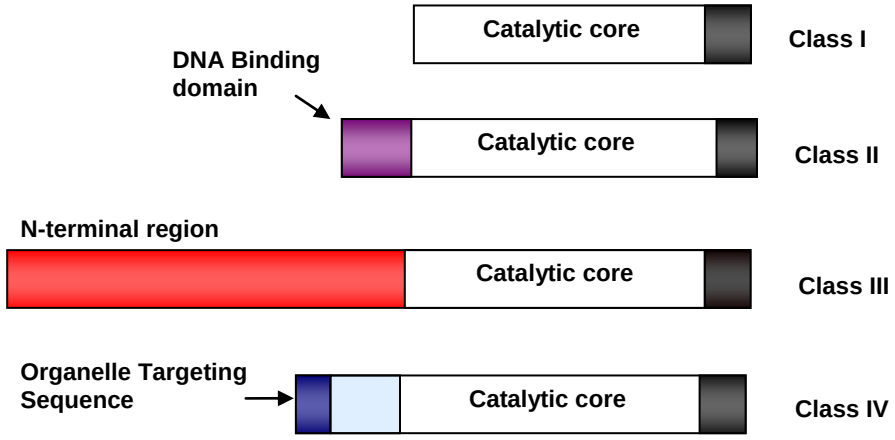


Figure 1.4 Diagrammatic primary structure comparison of the four classes of BPLs

1.4 Holocarboxylase synthetase (HCS)

Human BPL is commonly known as holocarboxylase synthetase (HCS) in the literature and will be referred to as HCS from here on. The coding sequence for HCS was first reported by Suzuki *et al.* (1994) using primers from the amino acid sequence of purified bovine BPL protein to identify cDNA clones, while Leon-Del-Rio *et al.* (1995) used functional complementation of an *E. coli birA* mutant to isolate the cDNA. The open reading frame of the HCS cDNA encodes a 726 residue protein, which Leon-Del-Rio *et al.* (1995) predicted to be M_r 80.8 kDa, which is monomeric in solution (Ingaramo and Beckett, 2009). The HCS

protein sequence contains a region homologous (amino acids 448-701) with BirA and this region is thought to be the biotin-binding region or the catalytic domain (Suzuki *et al.*, 2005). The HCS gene is comprised of 14 exons and is approximately 240 kb in length. Methionine 1 is located in exon 6 and a stop codon is present in exon 14 (figure 1.5 (A)). Out of the 13 introns, intron 9 is the longest at 130 kb. Figure 1.5 (B) summarises the 3 types of HCS mRNA that have been characterised. The first type starts at exon 1 and was isolated from human liver, type 2 starts at exon 3 and was isolated from lymphocytes and type 3 starts at exon 2 and was isolated from KG-1 myeloid cells (Yang *et al.*, 2001). Type 1 and 3 mRNA use 135 bp of exon 3 and type 2 mRNA uses 389 bp of exon 3. Exon 3 contains three stop codons in all three reading frames and therefore the initiation codon is the same for the three types of mRNAs. They found that the variations at the 5´ end of the mRNA did not change the amino acid sequence of HCS. Regardless of the presence of the three types of mRNA, Met1 starts at exon 6 for all three. Exons 6 to 14 are identical in all three mRNAs and it is therefore unclear why the three forms of mRNA are transcribed. Three alternative translation initiation sites of HCS have been proposed at Met1, Met7 and Met58 and utilisation of all three have been shown *in vivo* (Hiratsuka *et al.*, 1998). The initiation codons for Met1 and Met7 isoforms are both found within exon 6. The mechanism that determines alternative expression of these isoforms remains to be explored. The Met58 isoform is proposed to be expressed through differential splicing of mRNA, with the excision of exon 6 from the transcript and exon 7 providing the initiation codon (Yang *et al.*, 2001).

HCS splice variants have been identified in various cell types; which suggests cell specific regulation of mRNA processing. An analysis of the products of alternative initiation of translation found that 85.7 % were associated with cellular localisation (Cai *et al.*, 2006).

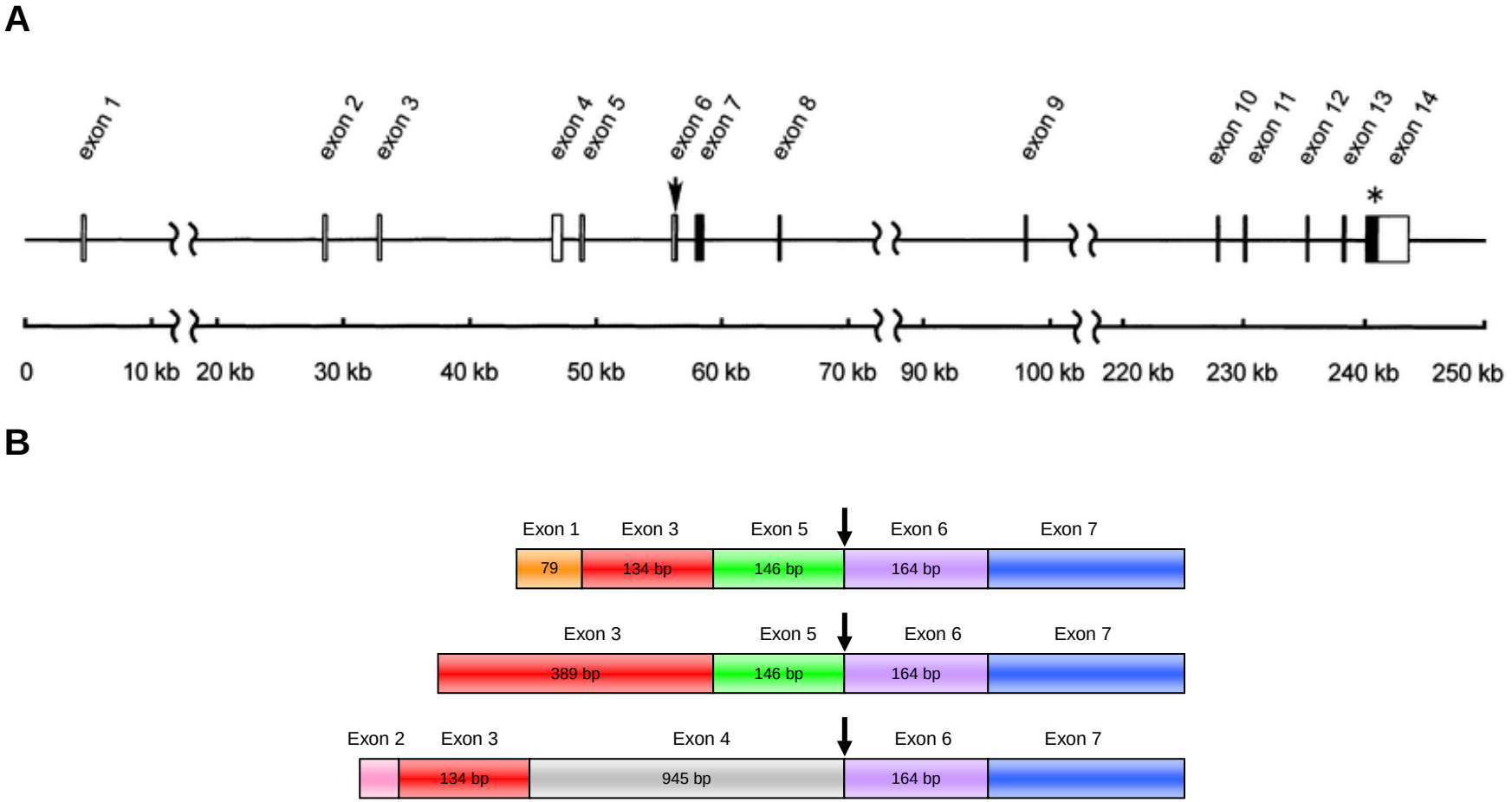


Figure 1.5 **(A)** Schematic structure of HCS gene. Coding regions are indicated by filled boxes or bars. The arrow and asterisk indicate the position of the Met1 and a stop codon. **(B)** Three types of 5´ region of HCS mRNA. Type 1 mRNA was isolated from human liver, type 2 mRNA was isolated from lymphocyte cDNA and type 3 mRNA was isolated from KG-1 myeloid cell line (Adapted from Yang *et al.*, 2001).

This suggests that differential initiation codon usage in HCS may direct subcellular localisation. The majority of the N-terminal truncated forms had significantly lowered isoelectric points, which potentially targeted them to different cellular compartments. Narang *et al.* (2004) reported a 66 kDa HCS band in both nuclear and cytosolic fractions. Fractionation studies on mouse 3T3-L1 cells found that over 70 % of all HCS activity localized to the cytosolic fraction (Chang and Cohen, 1983). Cohen *et al.* (1985) studied the subcellular distribution of HCS in rat liver and found 18 % of HCS activity was in the nuclear fraction, 18 % in the mitochondrial fraction, 5 % in the microsomal fraction and the majority (59 %) in the cytoplasm.

The significance of nuclear HCS is that it has been implicated in biotinylating histones *in vitro* (Hymes *et al.*, 1995; Stanley *et al.*, 2001; Narang *et al.*, 2004). This is despite the lack of any recognisable biotinylation sequence within histones. The detection of biotin on histones was performed by use of streptavidin, which Bailey *et al* (2008) showed was an artefactual detection of biotin, as streptavidin bound to the histones non-specifically. Healy *et al.* (2009) showed that histone H2A can be biotinylated non-enzymatically as incubation of recombinant H2A with chemically synthesized biotinyl-5´-AMP in the absence of HCS was labeled with biotin. In another study Healy *et al.* (2009) did not detect biotin attachment on native histones to a sensitivity of at least one part per 100 000, which suggested that the regulatory impact of biotin on gene expression was via alternate mechanisms.

Bailey *et al.* (2010) found that the three isoforms of HCS are predominantly localised in the cytoplasm and mitochondria and not in the nucleus, which is consistent with the distribution of HCS activity. The mechanisms responsible for mitochondrial activity have not been determined, especially since HCS does not have a mitochondrial localisation sequence. Bailey *et al.* (2010) found that unlike the longer isoforms, Met58 was detected in the nucleus, which suggests shuttling activity between the nucleus and the cytoplasm. Kinetic

analysis on the isoforms found a 40 % reduction in activity in the Met58 isoform as compared to the Met1 isoform (Sakamoto *et al.*, 1999), this indicated the importance of the N-terminal region of HCS, which has been implicated in catalysis. Ingaramo and Beckett (2009) found that Met1 HCS associates with model substrate p67 (propionyl-CoA carboxylase 67) two fold faster than Met58 HCS. Their finding indicated that residues 1-58 were important in substrate binding. In this thesis, full length or Met1 HCS was used, as it is the most active of the isoforms.

The X-ray structure for HCS has not been determined and therefore structure function studies have been performed to elucidate some information. Campeau and Gravel (2001) performed a series of N-terminal and C-terminal truncations of HCS in order to assess the ability of each truncation to catalyse biotinylation *in vivo* in *E. coli* (figure 1.6 (A)). This was performed by generating N-terminal deletions by exonuclease treatment and truncations were subcloned into reading frame variants of the expression vector. The C-terminal constructs were generated using PCR and restriction sites were inserted for subcloning in the appropriate vector. Each construct was assayed by determining its ability to complement *E. coli bir*A104, which confers a temperature-sensitive phenotype to BirA, at the restrictive temperature of 42°C.

Campeau and Gravel (2001) found varying results within a series of truncations of the N-terminal region. Truncations up to L166 demonstrated no distinguishable effect on the ability of HCS to biotinylate substrates or complement a conditional lethal BirA strain of *E. coli*. Truncations between L166 and T266 either did not complement or complemented weakly. Further truncations between T266 to K378 restored HCS function as growth was observed.

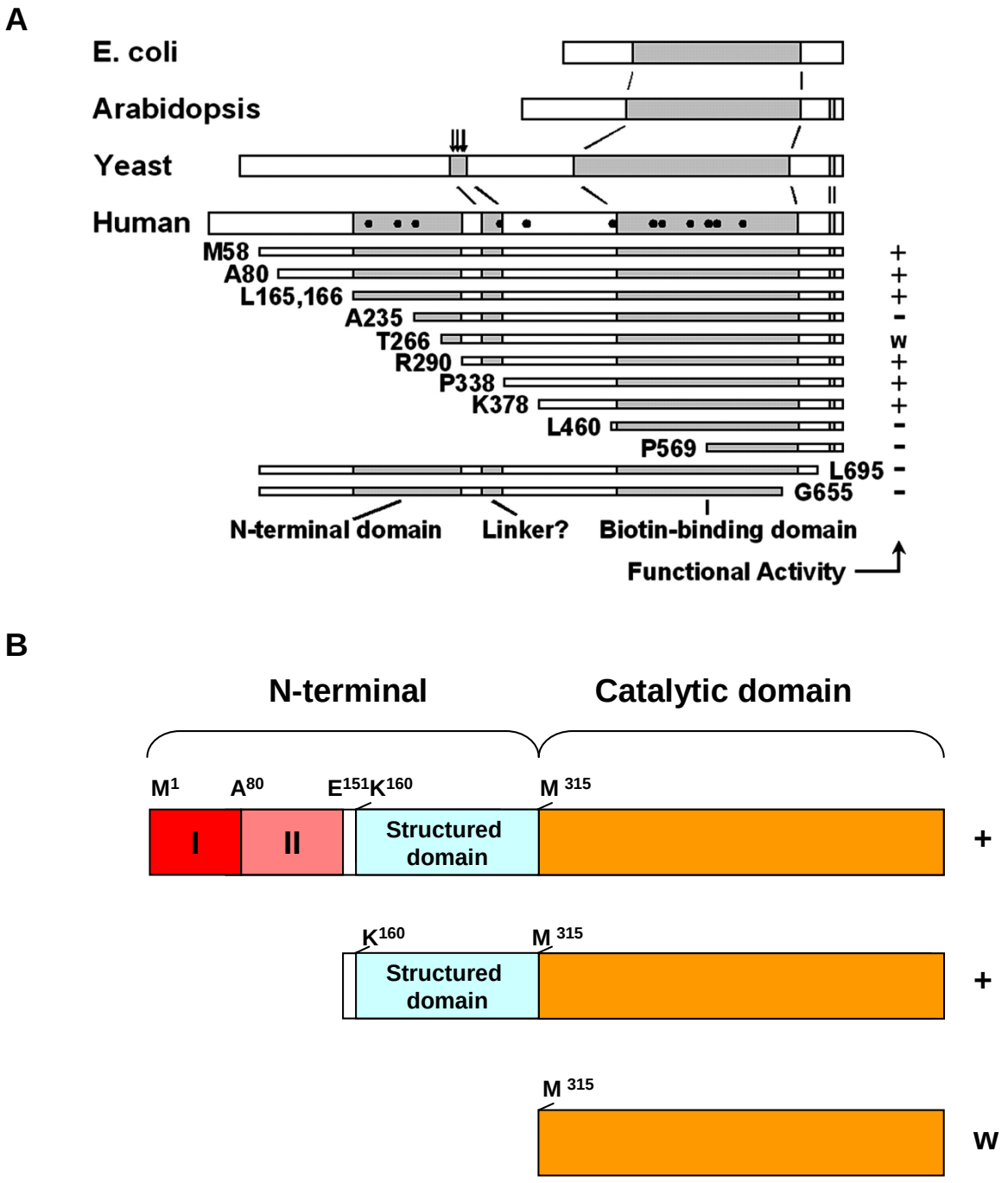


Figure 1.6 **(A)** Schematic of domain configurations in representative biotin protein ligases, including distribution of point mutations in HCS deficiency and summary of functional outcomes in deletion constructs of human HCS (Campeau and Gravel, 2001).
(B) Domain mapping of HCS by limited proteolysis (Swift, 2004). The N-terminal region contains three domains separated by protease sensitive linker regions. I (Met^{1} – Ala^{80}), II (Ala^{80} – Glu^{151}) and proposed structured domain or NTC (Lys^{160} – Arg^{314}). Activity is denoted +, no activity is – and weak activity is w.

Campeau and Gravel (2001) postulated that the region of the N-terminal domain between L166 and T266 is a structured domain that functions in the biotinylation reaction or may be involved in substrate recognition and that the N-terminal region may interact with the C-terminal region. The removal of as few as 37 amino acid residues from the C-terminal domain affects HCS activity by lowering the overall stability of the protein. A major limitation of this experiment was that quantitative enzyme kinetics to dissect these interactions were not performed.

Previous work in our laboratory using limited proteolysis has shown that the N-terminal region contains 3 distinct regions, one of which is a proposed structural domain containing residues G159-R314 (figure 1.6 B). These regions include M1-A80 (I), A80-E151 (II) and K160-D311 (structured domain) (Swift, 2004). This observation correlates with work performed by Campeau and Gravel (2001), with the proposed structure domain corresponding to the domain (L166-R290) that they postulated to interact with the catalytic domain. Based on Swift (2004) and using Met315 as the start site for the C-terminal domain, a proposed schematic diagram of domain structure of HCS was performed (figure 1.7). The domain boundaries of this proposed structure were used throughout this study.

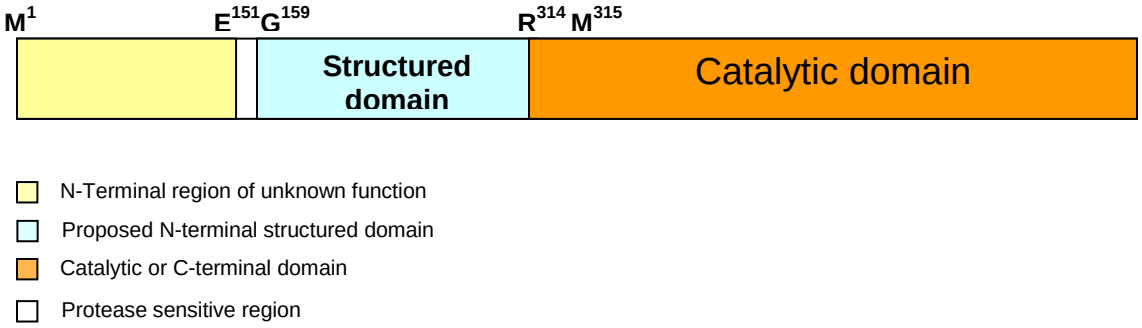


Figure 1.7 Schematic diagram of proposed HCS domain structure

Campeau and Gravel (2001) also used the N-terminally truncated HCS constructs, Met58, Ala80, Arg290, Pro338, and Lys378 that were able to complement *E. coli bir*A104 at 42°C to determine their substrate specificity. Each HCS construct was partially purified using Ni-NTA agarose and was tested for biotinylation of p67 (human propionyl-CoA carboxylase 67) or bacterial BCCP87 using a ^{14}C biotin incorporation assay. Met58 and Ala80 HCS biotinylationed both BCCP87 and p67. Arg290 and Pro338 HCS did not biotinylate BCCP87 but did biotinylate p67. Lys378 HCS showed preference for p67 and weak biotinylation of BCCP87. This was the first study to propose that the N-terminal region of HCS possesses "proof reading activity" that allows it to differentiate between various substrates.

1.5 Microbial BPLs provide structural and functional information on HCS

1.5.1 Class I BPLs

1.5.1.1 *Pyrococcus horikoshii* OT3 BPL (*Ph*BPL)

The X-ray structure for HCS has not been determined; however Class I and Class II BPL structures have been solved. The crystal structure of the BPL from *P. horikoshii*, a hyper-thermophilic archaebacterium, has also been solved by Bagautdinov *et al.* (2005). Class I *Ph*BPL consists of 2 domains (figure 1.8) and does not have the DNA binding domain of class II enzymes. The N-terminal domain (residues 1-187) is larger than the C-terminal domain (residues 191-235) and it contains a seven-stranded mixed β-sheet on one side and five α-helices on the other side. An unstructured loop links the two domains by connecting helix α_5 and strand β_8 (residues 187-190). The C-terminal domain consists of a five-stranded antiparallel β-sheet. The oligomeric state of *Ph*BPL in solution was reported as a constitutive dimer (figure 1.8) through the use of a dynamic light-scattering experiment and the two subunits of the homodimer are linked with a pseudo-2-fold symmetric interface

with hydrogen bonded β_1 strands to form an inter-subunit β-sheet (Bagautdinov *et al.*, 2005).

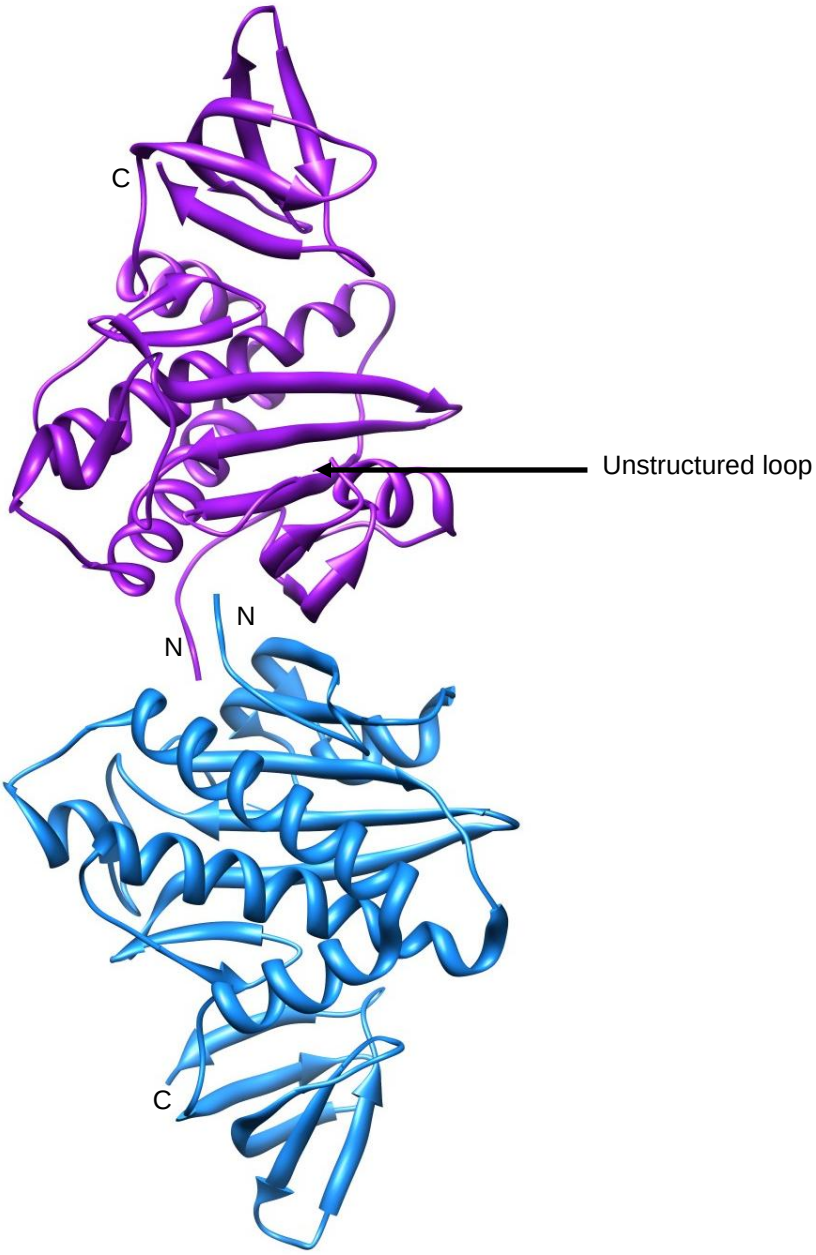


Figure 1.8 The X-ray crystal structure of apo *Ph*BPL homodimer. The subunits are distinguished by colouring (Bagautdinov *et al.*, 2005). Figure was generated by UCSF Chimera using PBDID 1WQ7. N is the N terminus and C is the C terminus. Unstructured loop becomes structured upon ligand binding.

The X-ray crystal structure of *Ph*BPL was consistent with the gel filtration data in that it is a constitutive dimer in solution. Bagautdinov *et al.* (2005) studied the conformational changes in the structure of *Ph*BPL that occur after binding of biotin, ADP and biotinyl-5´-AMP. This was the first study that provided information on the structural conformation of a BPL when each of the three ligands was bound. Upon ligand binding, the overall backbone structures of the complexes are very similar (figure 1.9 (A-C)). Each ligand binds to its pocket and induces minor changes in local conformations. The main difference between unliganded and liganded *Ph*BPL are localised to the N-terminal surface loops. When biotin is bound to *Ph*BPL it interacts with the strands β_2, β_3, β_6 and β_7, β_2-β_3 loop and the N-termini of helices α_1 and α_2 (figure 1.9 (B)). The biotin carboxyl group and ureido nitrogen atoms interact by hydrogen bonding with the hydrophilic pocket. The hydrophobic tail and thiophene ring interact with the hydrophobic wall. The glycine residues located at the bottom of the pocket accommodate the ureido and thiophene heterocycles of biotin. The determination of the structure of *Ph*BPL in complex with ADP by Bagautdinov *et al.* (2005) was the first report of a BPL with a nucleotide bound in the appropriate pocket. They found the position of adenine in the structure of *Ph*BPL liganded with ADP (figure 1.9 (C)). ADP binds to the site adjacent to the biotin binding pocket and is held in place by a network of electrostatic and polar interactions with part of the β_2-β_3 loop, the β_6 strand, the β_7-α_3 loop and the strand β_{12}. Figure 1.9(D) shows the structure of *Ph*BPL in complex with biotinyl-5´-AMP and it was the first solved structure of a BPL with the reaction intermediate. Here biotinyl-5´-AMP adopts an unexpected v-shaped geometry that occupies both the ATP and biotin pockets.

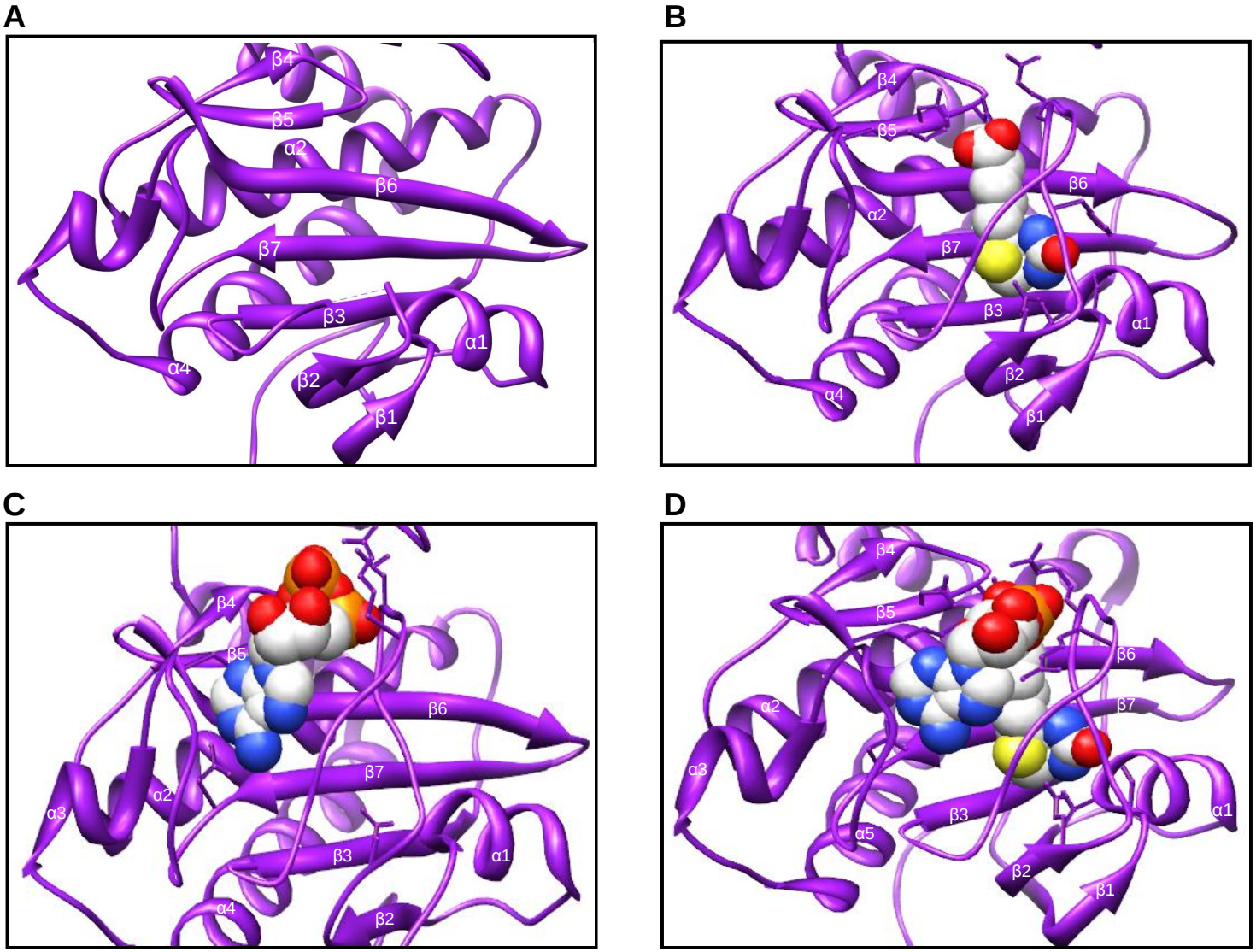


Figure 1.9 A close-up view of the *Ph*BPL active site in **(A)** the unliganded form, and in the liganded forms with **(B)** biotin, **(C)** ADP and **(D)** biotinyl-5′-AMP (Bagautdinov *et al.*, 2005). Figure was generated by UCSF Chimera using PBDIDs 1WQ7, 1WPY, 1WNL and 1WQW, for A-D, respectively. Atoms in each ligand are viewed as spheres.

Bagautdinov *et al.* (2008) solved the crystal structure of PhBPL in complex with the biotin carboxyl carrier protein (*Ph*BCCP) (figure 1.10). This is the only example of such a complex solved so far. Figure 1.10 shows the orientations of bound biotin and adenosine. The *Ph*BPL-*Ph*BCCP complex involves the formation of a large intermolecular β-sheet, which is solvent-exposed on one side to accommodate the biotinyl-5´-AMP. The main holes of the complex are occupied by biotin and adenosine. There is also a degree of plasticity around the active site of *Ph*BPL to accommodate its substrate *Ph*BCCP. There is especially conformational flexibility of the C-terminal domain loop (Ile226–Asp229) in the side hole to the active site provides an entry for the *Ph*BCCP lysine residue and an exit pathway for the product biocytin residue of holo-BCCP. As figure 1.10 (B) shows, only the the β-turn loop containing Lys115 is buried in the *Ph*BPL side whole, the rest of the *Ph*BCCP is solvent exposed. In this thesis, the interaction between human BPL and its substrate is characterized. Due to the fact that the structure of human BPL is yet to be solved, the Bagautdinov *et al.* (2008) study is useful in understanding this interaction at the structural level.

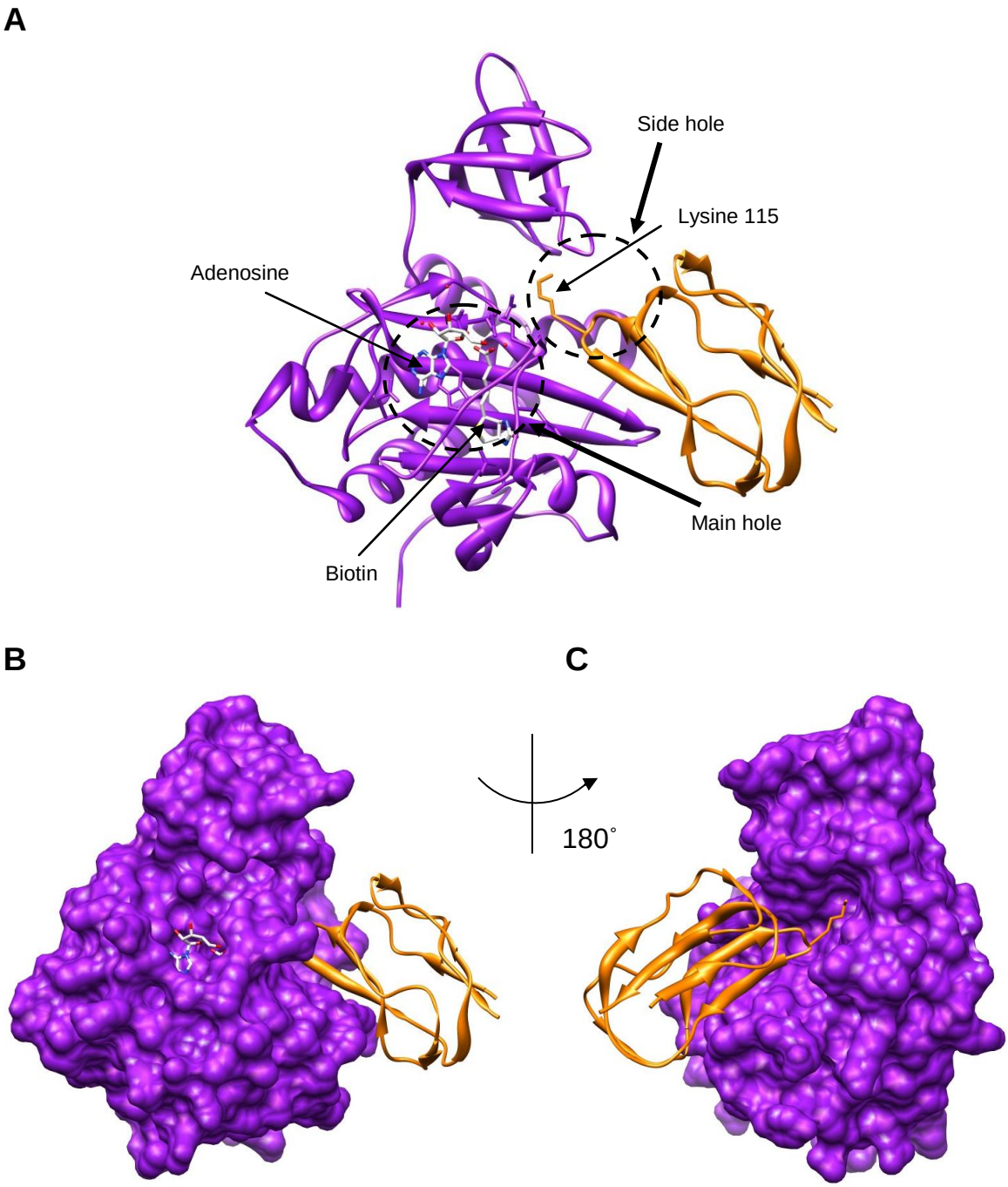


Figure 1.10 The *Ph*BPL (purple) in complex with *Ph*BCCP (orange) (Adapted from Bagautdinov *et al.*, 2008). **(A)** *Ph*BPL presented in ribbon format, **(B)** is *Ph*BPL with the surface shown and (C) PhBPL back view with surface displayed. Figure was generated by UCSF Chimera using PBDID 2EJF. Each ligand is viewed in stick.

1.5.2 Class II BPLs

1.5.2.1 BirA

The best characterised BPL Class II enzyme is *E. coli* BPL, also known as the biotin inducible repressor A (BirA). BirA is a bifunctional protein as it catalyses the attachment of biotin onto the biotin carrier protein (BCCP) and acts as a transcriptional repressor of the biotin biosynthetic operon. In the presence of apoBCCP, BirA acts as a ligase. When the enzyme has been biotinylated and the cellular demand for biotin is low, BirA acts as a transcriptional repressor to prevent the expression of enzymes involved in biotin biosynthesis. BirA is a 35.5 kDa protein that has been crystallised and its structure determined to 2.3 Å. This revealed that the enzyme is composed of three distinct domains (figure 1.11). The domains include an N-terminal DNA binding domain, a central catalytic domain and a C-terminal domain of unknown function (Chapman-Smith *et al.*, 1999). Therefore BirA is a truly multi-functional protein and key regulator of all biotin related events in *E. coli*.

The N-terminal domain consists of a helix-turn-helix motif characteristic for DNA binding. The C-terminal domain functions as the catalytic domain (Chapman-Smith *et al.*, 1999; (Brown and Beckett, 2005). Enzyme kinetic data has shown that binding of biotin and ATP to BirA occurs in an ordered manner, with biotin binding first (Xu and Beckett, 1994). The residues that contact biotin in the crystal structure are highly conserved (Wilson *et al.*, 1992; Tissot *et al.*, 1997). Within this domain there are several loops that are disordered in the crystal structure but which become more ordered when in complex with biotin or biotinyl-5´-AMP. The loop that comprises residues 211-223 contains a subtilisin-sensitive site C-terminal to S217 that is protected against proteolytic cleavage when biotinyl-5´-AMP is present. This indicates that there is a structural transition following ligand binding in the first catalytic step (Xu and Beckett, 1994).

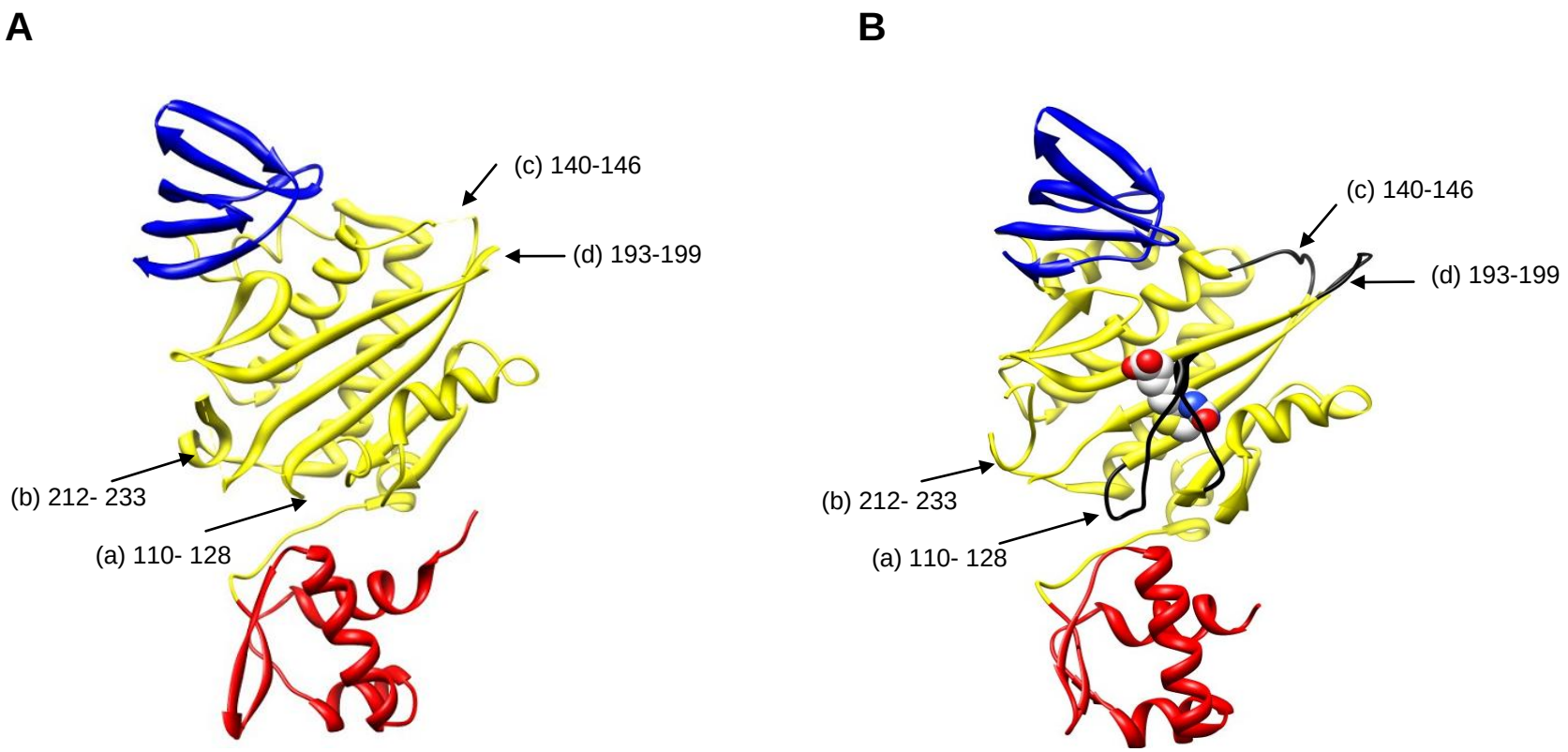


Figure 1.11 The apo **(A)** and holo **(B)** structures of BirA solved by X-ray crystallography are shown in ribbon presentation. The structure contains three domains; the N-terminal domain, depicted in red, the central or catalytic domain in yellow and the C-terminal cap in blue. **(A)** Several surface loops, indicated by arrows with residue range, are assumed to be mobile. Figure generated by UCSF Chimera using PBD 1BIA (Wilson *et al.*, 1992). **(B)** The holo structure of BirA was solved by co-crystallising BirA with biotin (viewed in sphere). The surface loops that were disorded and became ordered when biotin is bound are shown in black. Loop (a) is involved in ligand binding, loops (c) and (d) form part of the dimerisation interface and (b) remained disorded (Wilson *et al.*, 1992). Figure was generated by UCSF Chimera using PBDID 1HXD.

1.5.3 Class III BPLs

1.5.3.1 Yeast biotin protein ligase (yBPL)

Most eukaryotic BPLs have been found to be around twice the size of their bacterial counterparts. The BPL of *Saccharomyces cerevisiae* is composed of 690 amino acids and is predicted to be 76.4 kDa in mass. The C-terminal region of eukaryotic BPLs has been found to show conservation with that of BirA, suggesting that this site contains the catalytic site (Cronan, 1995; Polyak *et al.*, 1999). The similarity is clear in residues that are in contact with biotin in the crystal structure of *E. coli* BirA and those associated with mutations in BirA that result in an increase in K_M for biotin are highly conserved (Wilson *et al.*, 1992; Cronan, 1995; Chapman-Smith *et al.*, 1999). This is consistent with the fact that unlike bacteria, both yeast and humans require an exogenous source of biotin and do not bind DNA to regulate biotin biosynthesis (Chapman-Smith *et al.*, 1999; Polyak *et al.*, 1999).

Polyak *et al.* (1999) performed limited proteolysis studies on the yBPL revealing a flexible linker that separates the N-terminal 27 kDa domain and the C-terminal 50 kDa domain. Addition of the ligands ATP and biotin decreased the susceptibility of yBPL to proteolysis, suggesting that substrate binding lead to a conformational change in the structure thereby preventing lysis of the linker region. Polyak *et al.* (1999) also showed that the N-terminal domain of yBPL is required for activity as the expression of N-terminally truncated variants of the enzyme did not complement *E. coli* strains that were defective in BPL activity. They also performed *in vitro* assays with purified N-terminally truncated yBPL and found that the removal of the N-terminal domain lowered the enzyme's activity by greater than 3,500-fold.

1.6 Comparison of the N-terminal and C-terminal regions of HCS (and other Eukaryotic BPLs) with N- and C- terminal regions of microbial BPLs

Eukaryotic BPLs have multiple substrates within the cell, unlike its bacterial counterpart *E. coli* which has only one protein substrate. It is not known whether there is preferential biotinylation of substrates in Eukaryotes. Although the BPL recognises the biotin domain within each enzyme, it is unknown whether substrate specificity between enzymes occurs. The catalytic core of BPLs is conserved across kingdoms (figure 1.4 and figure 1.12), this suggests a common catalytic mechanism conserved through evolution. There is significant sequence variation in the N-terminal regions of BPLs across kingdoms (figure 1.4 and figure 1.13). HCS contains a large N-terminal extension compared to microbial BPLs and bears no sequence homology. The large N-terminal extensions of yeast and mammalian BPLs are not believed to be involved in DNA binding (as for BirA) as no DNA binding domain is present. The function of this N-terminal extension is currently unknown, although it has been implicated in the catalytic function of HCS (Campeau and Gravel, 2001) and yeast BPL (Polyak *et al.*, 1999).

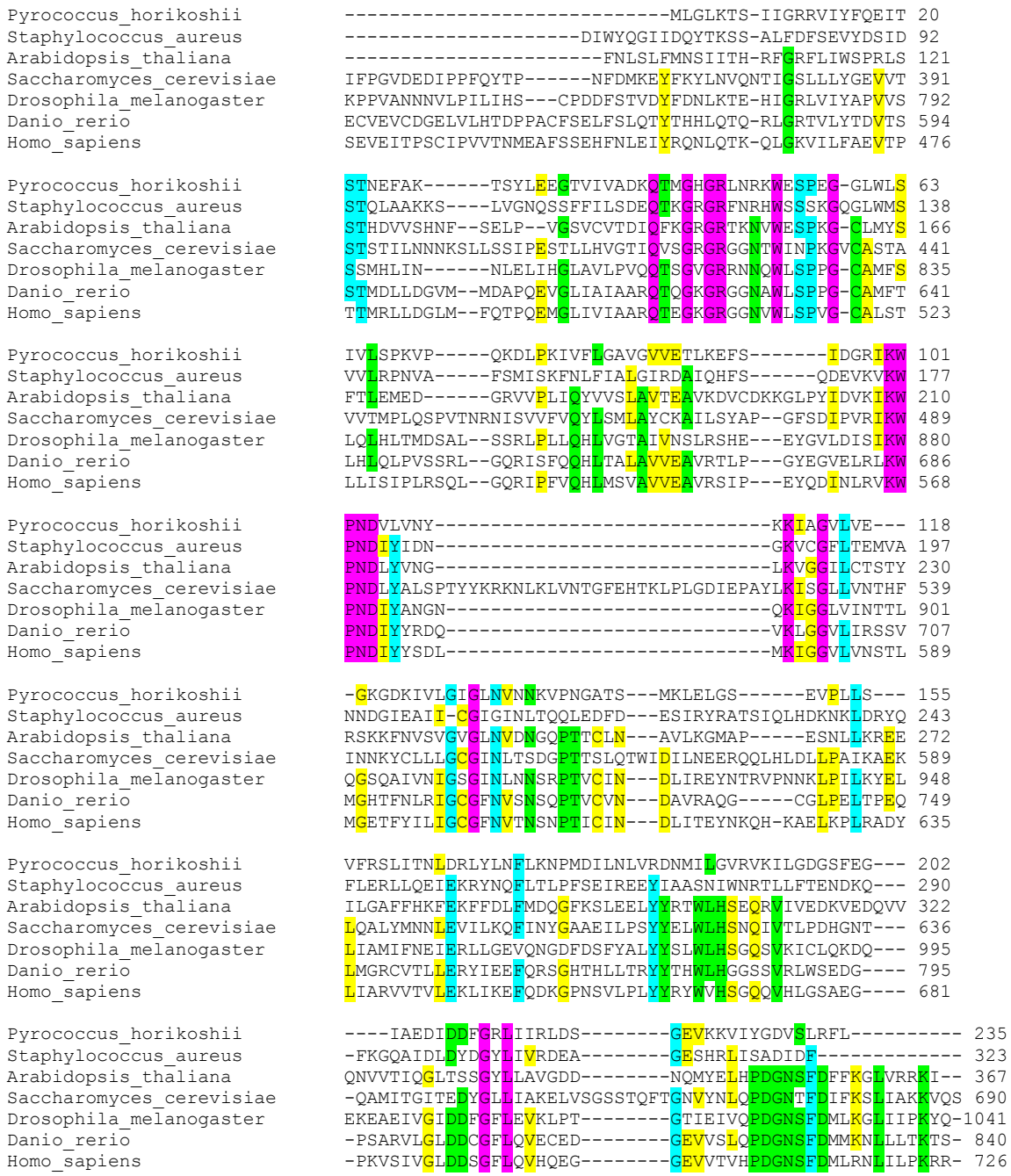


Figure 1.12 Conservation of the catalytic core or C-terminal region of BPL through evolution.
A multiple sequence alignment was performed in ClustalW using *P. horikoshii*, *E.coli*, *S. aureus*, *Arabidopsis thaliana*, *S. cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Danio rerio*, *Mus musculus* and *Homo sapiens*. Colours indicate the conservation of amino acids; 100 %, 85 %, 70 % and 60 % identity.

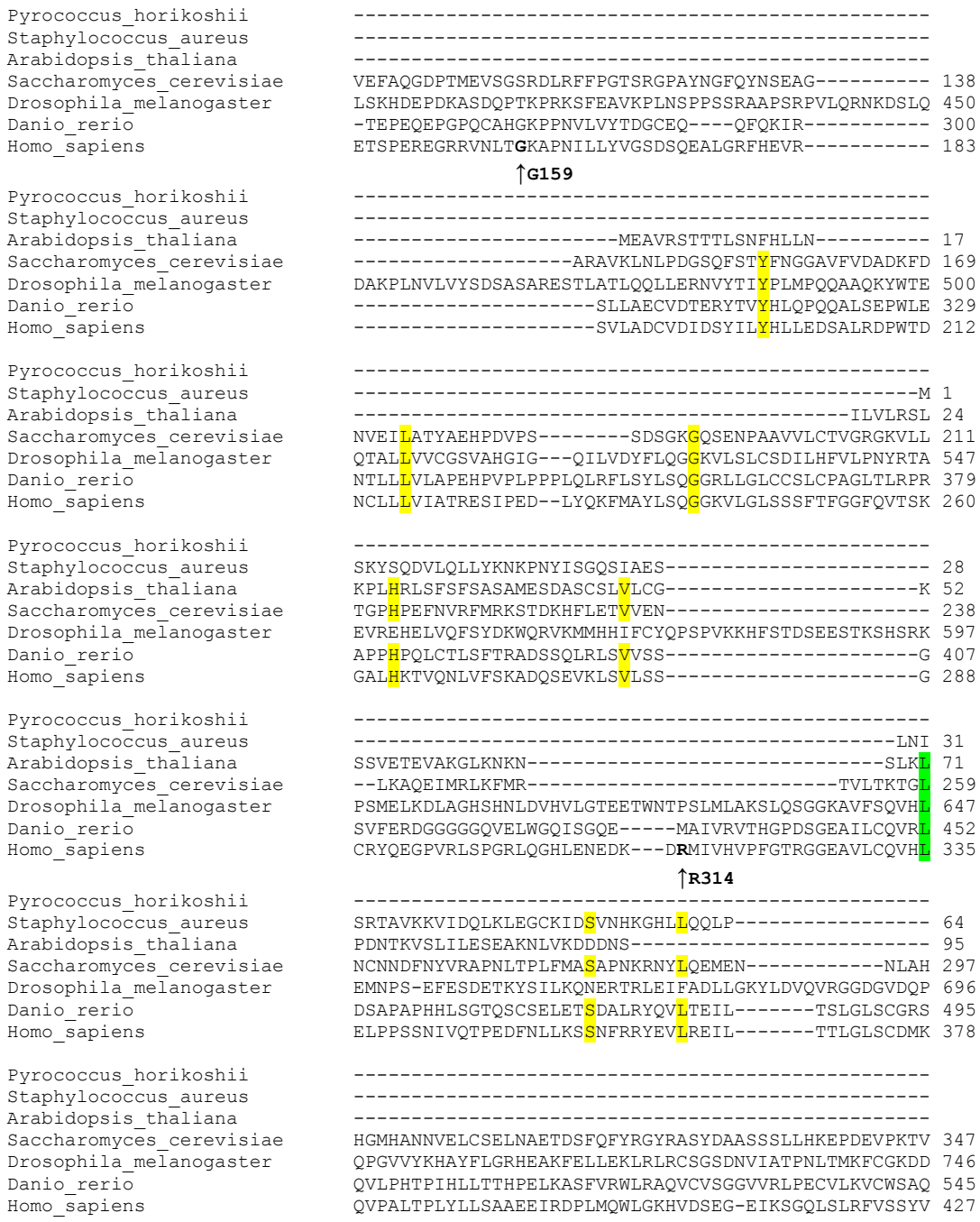


Figure 1.13 Comparison of the N-terminal region of BPL through evolution.
A multiple sequence alignment was performed in ClustalW using *P. horikoshii*, *E. coli*, *S. aureus*, *Arabidopsis thaliana*, *S. cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Danio rerio*, *Mus musculus* and *Homo sapiens*. Colours indicate the conservation of amino acids; ▮ 70 % and ▮ 60 % identity. The proposed N-terminal structured domain of HCS, G159-R314, is displayed in bold.

1.7 Multiple Carboxylase Deficiency (MCD)

All five biotin-dependent enzymes (figure 1.1) are inactivated when there are defects in biotin metabolism. This leads to the severe condition known as multiple carboxylase deficiency (MCD) (Feldman *et al.*, 1981; Narisawa *et al.*, 1982; Saunders *et al.*, 1979; Wolf and Feldman, 1982). MCD arises due to a defect in the biotinylation of carboxylases (HCS deficiency, OMIM 253270), inefficiency in recycling biotin in the cell (biotinidase deficiency, OMIM 253260) or a failure in biotin transport. Biotin deficiency results from either inadequate dietary intake or by congenital defects in biotin utilisation, recycling or intestinal adsorption (Baumgartner and Suormala, 1997; Wolf, 1995). Consequently, MCD can occur at various ages (from birth to the teenage years) and with a wide range of symptoms including ketoacidosis, lactic acidosis, feeding difficulties, neurological abnormalities (including subependymal cysts, hypotonia, seizures and ataxia) and rashes. In severe cases, or if left untreated, MCD can lead to coma or death (Baumgartner and Suormala, 1997; Wilson *et al.*, 2005).

As HCS is solely responsible for the biotinylation of the five biotin-dependent enzymes. Patients with inherited HCS-deficiency (OMIM 253270) present the same biochemical and clinical symptoms as for MCD (Baumgartner and Suormala, 1997; Wolf, 1995). HCS-deficiency is a severe juvenile condition that occurs from a few days to several years post birth. The age of onset and severity of the disease corresponds with the effect of mutation on the HCS enzyme (Aoki *et al.*, 1999; Suormala *et al.*, 1998). Mutations in the HCS gene that give rise to MCD can be classified in two groups. The first group is of mutations that give rise to MCD that localise at the catalytic site (figure 1.14) and within exons 9-14 (figure 1.5 A). A molecular explanation for the efficacy of biotin treatment has been provided from biochemical studies on HCS, which showed an increased K_M for biotin (Aoki *et al.*, 1999; Dupuis *et al.*, 1999; Suzuki *et al.*, 2005; Sweetman and Nyhan, 1986; Yang *et al.*, 2001). This group of patients respond to treatment with biotin supplementation

(10-20 mg/day), which can provide sufficient substrate to increase the enzymatic function of HCS and abolish the biochemical and clinical symptoms of MCD (Wolf *et al.*, 1981; Zempleni and Mock, 2000). An increase in biotin serum levels due to the therapy reverses the effect of the mutations.

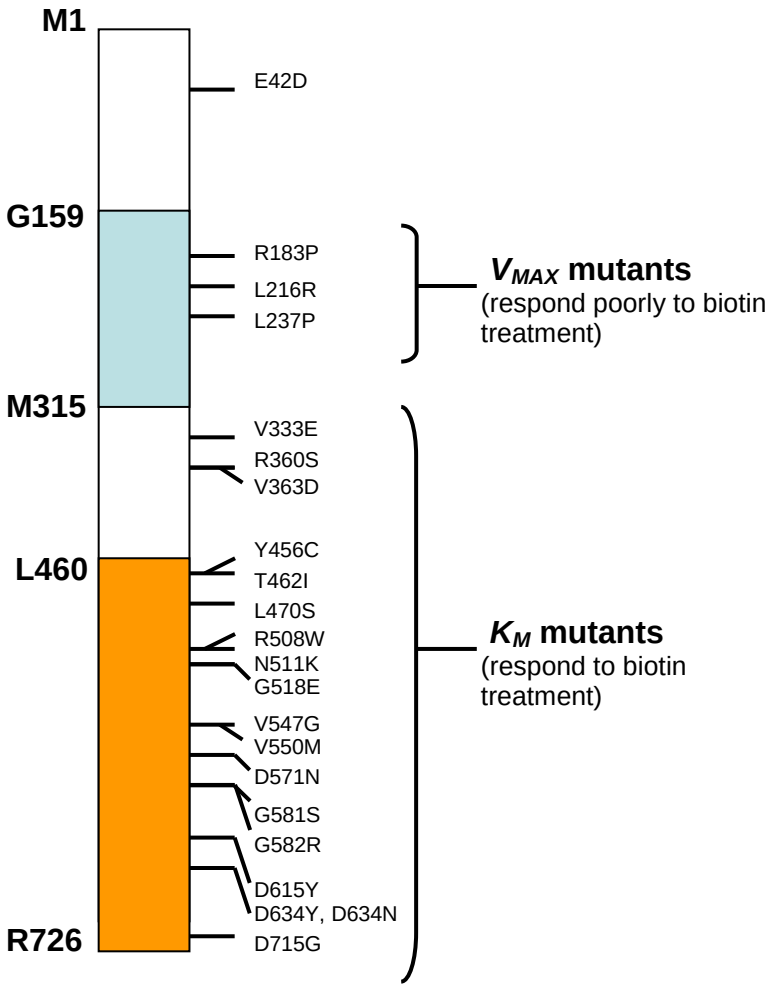


Figure 1.14 Schematic representation of the HCS and the mutations that give rise to MCD. HCS protein, highlighting the conserved C-terminal catalytic region (orange box) and the domain in the N-terminal extension required for catalysis (light blue box). Positions of mutations arising in MCD are noted on the right of the representation and are grouped under either V_{MAX} and K_M mutants (Adapted from Suzuki *et al.*, 2005).

The other group of HCS-deficient patients do not respond to oral biotin therapy, even with doses as high as 200 mg. In some cases for these patients the outcome is coma or death

(Aoki *et al.*, 1995; Dupuis *et al.*, 1996; Morrone *et al.*, 2002; Wilson *et al.*, 2005; Dupuis *et al.*, 1999; Zempleni and Mock, 2000; Ames *et al.*, 2002). The HCS mutants that give rise to MCD in this group of patients contain mutations in the N-terminal domain of HCS (figure 1.14) and exon 7 (figure 1.5 A). These mutations were found to have a significant decrease in V_{MAX}, however, a similar K_M for biotin was reported (Aoki *et al.*, 1997).

Table 1.1 is a summary of the known effects on the K_M for biotin and V_{MAX} for MCD mutations that have been characterised *in vitro*. It is clear that for mutations in the catalytic region of HCS (residues 336-726) the K_M for biotin is elevated 3 to 44 fold compared to the wild type enzyme.

Table 1.1 Effect of MCD mutants on HCS activity. HCS activity was measured by expression of mutant HCS proteins in HCS deficient fibroblasts using apo-BCCP as a substrate except where indicated.

Mutation	Activity (% wild type)	K_M biotin (fold wild type)	References
Wild type	100	1	(Aoki *et al.*, 1999)
E42D	120	ND	(Yang *et al.*, 2001)
R183P*	1.7	0.6	(Sakamoto *et al.*, 1999)
L216R*	0.3	1.4	(Sakamoto *et al.*, 1999)
L237P*	1.2-4.3	0.4-1.2	(Sakamoto *et al.*, 1999, Aoki *et al.*, 1997)
V333E	2-10	1.5	(Sakamoto *et al.*, 1999, Aoki *et al.*, 1997)
R360S	22	ND	(Yang *et al.*, 2001)
V363D	3.7	1.1	(Sakamoto *et al.*, 1999)
Y456C	0.2	ND	(Yang *et al.*, 2001)
T462I	<10	ND	(Aoki *et al.*, 1999)
L470S	4.3	ND	(Yang *et al.*, 2001)
R508W	34.5**	23**	(Burri *et al.*, 1985)
V547G	3.4	ND	(Yang *et al.*, 2001)
V550M	16.6	6.5	(Aoki *et al.*, 1997)
D571N	0.1	ND	(Aoki *et al.*, 1997)
G581S	<10	44.3	(Sakamoto *et al.*, 1999, Aoki *et al.*, 1999)
delT610	14	2.9	(Sakamoto *et al.*, 1999, Aoki *et al.*, 1999)
D634Y	12	ND	(Yang *et al.*, 2001)
Del(c.C2279)	0.1	ND	(Aoki *et al.*, 1999)
Del(c.T1876)	0.1	ND	(Aoki *et al.*, 1999)

* N-terminal HCS mutants
** HCS activity measured in patient fibroblasts using apoPCC from rat liver as substrate
ND is not determined

1.7.1 K_M MCD mutants

There are several reports of MCD patients who have mutations in the C-terminal region of HCS and show an elevated K_M for biotin. There is a large variation in the K_M values obtained due to the different apo-carboxylase substrates used (Burri *et al.*, 1981; Burri *et al.*, 1985; Suzuki *et al.*, 1996; Aoki *et al.*, 1997; Dupuis *et al.*, 1999; Sakamoto *et al.* 1999; Sakamoto *et al.*, 2000; Morrone *et al.*, 2002). Some of mutations that affect K_M of biotin cluster around the biotin binding pocket as can be seen in the model of the C-terminal region of HCS (figure 1.15) which was generated based on the X-ray co-ordinates for *S. aureus* BPL using MODELLER v6.1 (Eswar *et al.*, 2006) package by Associate Professor Grant Booker (unpublished).

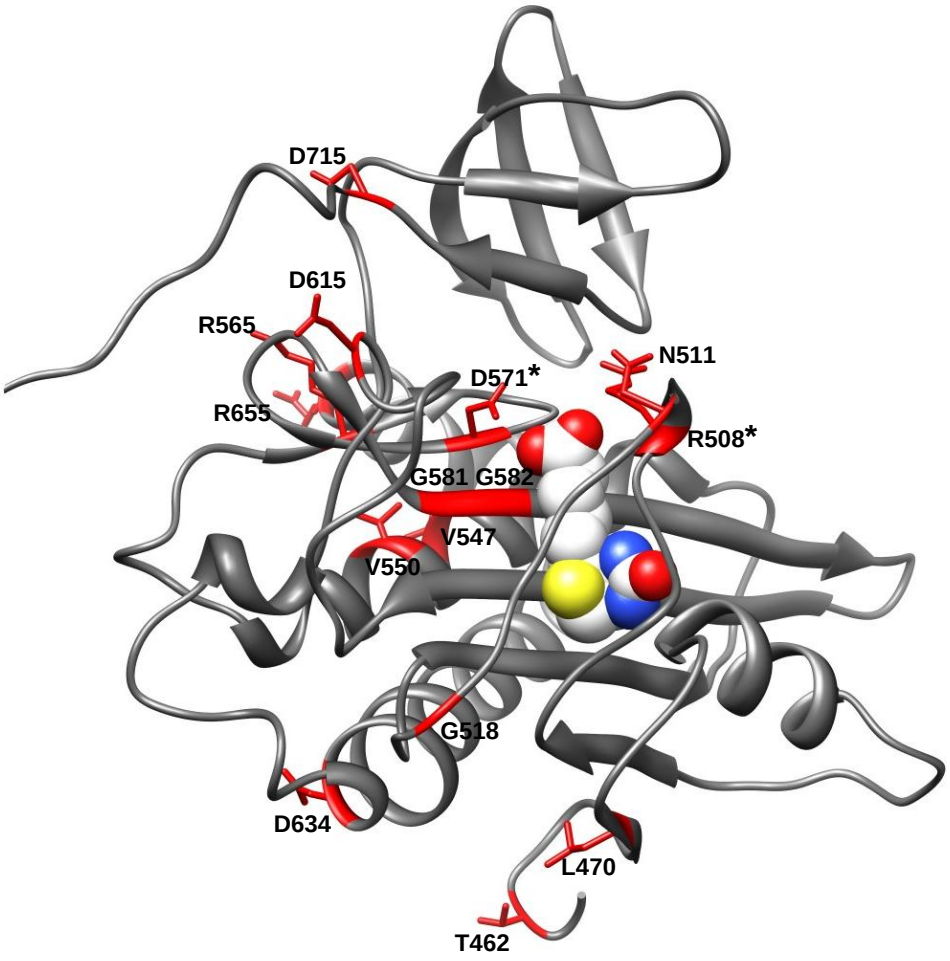


Figure 1.15 Mutations leading to Multiple Carboxylase Deficiency. Residues that are mutated in MCD have been highlighted (in red) on the C-terminal region of HCS model with biotin shown in the active site (in sphere). Figure was generated by UCSF Chimera. Asterics are the mutations studied in the literature.

The R508W HCS mutation is the most common mutation found in a variety of ethnic groups with MCD. This arginine is located at a highly conserved residue in all BPLs. In BirA Arg118 is located on a flexible loop (residues 110-128) that becomes structured upon ligand binding (Wilson *et al.*, 1992). Mutation of this residue results in promiscuous biotinylation caused by enhanced dissociation of the labile reactive intermediate biotinyl-5´-AMP from the BPL active site (Choi-Rhee *et al.*, 2004). The D571N HCS mutation occurs at a very highly conserved position within the KWPND motif present in all BPLs. This position, while it does not directly contact biotin or ATP, is critical for activity (Aoki *et al.*, 1999). Other mutations most likely result in subtle conformational changes that influence ligand binding. From the modelling it can be seen that these mutations in the catalytic domain increase the K_M for biotin. This can be overcome by increasing the concentration of biotin, thereby providing a molecular explanation for the efficacy of oral biotin therapy.

1.7.2 V_{MAX} MCD mutants

The V_{MAX} mutations have been mapped outside the catalytic site and within the N-terminal extension of HCS. The mutations in the N-terminal region were found to have a significant decrease in V_{MAX}, however, a similar K_M for biotin was reported (Aoki *et al.*, 1997). Patients with these mutations, namely R183P, L216R and L237P HCS, vary in their responsiveness to biotin treatment. There is no clinical data available on the one reported case containing the R183P mutation (Sakamoto *et al.*, 1998).

The L237P HCS mutation was the first N-terminal mutation of MCD reported (Aoki *et al.*, 1995) and recombinantly expressed and has severely reduced biotinylation activity *in vitro* (Aoki *et al.*, 1997; Sakamoto *et al.*, 1999). Responsiveness of patients with this mutation to biotin therapy varies widely due to genetic background. Table 1.2 summarises the studies conducted on the L237P HCS mutation. These cases suggest that the L237P mutant does

retain a low biotinylation activity, the effect of the L237P mutation was more severe when patients were heterozygous for L237P (with the other allele containing a truncated non-functional HCS).

Table 1.2 Treatment of MCD cases containing the L237P mutation

Patient	Genetic background	Treatment (Biotin)	Prognosis	References
1	Homozygous	40 mg/day	Normal at 9 months, lowered IQ at 5 years	(Aoki *et al.*, 1995)
2	Heterozygous	80 mg/day	Age 7, delays in mental development	(Aoki *et al.*, 1995)
UW	Homozygous	40 mg/day	Remained asymptomatic despite elevated excretion of 3-hydroxy-isovaleric acid in urine	(Sakamoto *et al.*, 1998)
1	Heterozygous	N/A	Died within the first 5 days post birth	(Yang *et al.*, 2001)
3	Heterozygous	N/A	Died within the first 5 days post birth	Yang *et al.*, 2001)

A study by Sakamoto *et al.* (1999) showed lymphoblasts and fibroblasts derived from MCD patients bearing the c.710T>C p.L237P mutation had severely compromised PCC activity (<10% wild type PCC). Biotin supplementation was only able to increase PCC activity to 20 % that of wild type at 100 µM biotin. This was in contrast to the biotin responsive HCS mutant c.1741G>A p.G581S, where PCC activity was completely restored at 100 µM biotin. These studies suggest that V_{MAX} is an important factor in the pathology of the disease, and there is a correlation between V_{MAX} and responsiveness to biotin therapy. No other similar studies were performed on the other biotin-dependent enzymes.

Like the L237P mutation, responsiveness of patients with the L216R HCS mutation to biotin therapy varies due to genetic background. Table 1.3 summarises the studies conducted on the L216R HCS mutation. The effect of the L216R mutation is more severe when patients were homozygous for L216R; this is in contrast to the findings with the L237P mutation.

Table 1.3 Treatment of MCD cases containing the L216R mutation

Patient	Genetic background	Treatment (Biotin)	Prognosis	References
VE	Heterozygous (L216R/V363D)	40 mg/day	Biotin responsiveness was due to the biotin responsive allele V363D	(Dupuis *et al.*, 1999)
1	Homozygous	100 mg/day	Symptoms evident within 12 hours post birth. Little response to biotin treatment even at high doses	(Morrone *et al.*, 2002)
1, 3-6	Homozygous	40 mg/day	Patients died between 3 days and 3 years. When biotin was not administered (2 cases) babies died within 7 days. Lower mean birth weight.*	(Wilson *et al.*, 2005)
2	Homozygous	20 mg/day	Showed slight improvement initially, by 18 months showed severe developmental delays. Lower mean birth weight.*	(Wilson *et al.*, 2005)
7	Homozygous	20 mg/day	Showed mild developmental delays and "relatively normal" skin. Mild lactic acidosis. Lower mean birth weight.*	(Wilson *et al.*, 2005)

* Lower mean birth weight of 2760 g compared to a mean of 3700 g of healthy babies. Unique to L216R mutant.

These studies show that the V_{MAX} mutants have severe symptoms and in some cases lead to death. Although this is a rare disease, it is still important to improve the quality of life of patients and to prevent the loss of life. To this end, suitable treatments for MCD patients with V_{MAX} mutations are required. In this thesis, the molecular mechanism of the V_{MAX} mutants is described and potential treatments for MCD are presented.

The L216R HCS mutant was further characterised by the Bailey *et al.* (2008) study where they investigated cell lines from two HCS-deficient patients homozygous for the c.647T>G L216R allele. The study found that the growth of the patients' fibroblasts was compromised compared with normal or healthy fibroblasts. The patients' fibroblasts were not sensitive to biotin-depletion from the media and growth rates were not restored by the re-administration of biotin. The L216R HCS mRNA was detected in the MCD cell lines. However, expressed protein was not detected in the cell lines. Enzyme activity was also not detected in the patients' cells. They performed *in vitro* assays on recombinantly expressed L216R and

determined that the enzyme activity was severely compromised and the activity was not increased by the addition of biotin. The study also showed that the turn-over rate (or rate of degradation) for L216R HCS was double that of wild type HCS and thereby it is less stable than the wild type. This data suggests that the decreased half-life of L216R HCS contributes to the incomplete responsiveness of the V_{MAX} MCD patients to biotin therapy.

The MCD mutations L216R and L237P HCS are outside of the catalytic site and map to the proposed N-terminal structured domain (figure 1.14). The N-terminal domain of HCS has been implicated in catalysis (Campeau and Gravel, 2001) by N-terminal truncation studies and the significance N-terminal MCD mutants indicate the importance of this region for the function of HCS. It has not been clearly shown how these N-terminal point mutations affect the activity of HCS and determining their effect will provide a more specific molecular mechanism for MCD.

This thesis will functionally and structurally characterise the N-terminal region of HCS and will also provide a molecular explanation for the V_{MAX} MCD mutants as this has not been adequately resolved. It has been proposed that the N-terminal extension bearing these mutations is involved in the reaction mechanism, possibly by compromising the protein:protein interaction with the biotin-dependent enzyme substrates (Campeau and Gravel, 2001; Hassan *et al.*, 2009; Lee *et al.*, 2009; Lee *et al.*, 2010). Here, new data is presented to demonstrate the function of the N-terminal extension in catalysis and in turn provide a novel molecular mechanism to explain the basis of HCS-deficiency.

It is our goal to understand the role of the N-terminal domain of HCS in catalysis. To achieve this I will address two hypotheses:

The first hypothesis is the proposed structured N-terminal domain interacts with the catalytic/C-terminal domain.

The second hypothesis is the first partial reaction of HCS proceeds in the V_{MAX} MCD mutants, but the second partial reaction is affected by the mutations.

To address these hypotheses the aims of this study were to:

1. Provide evidence for the molecular interaction between the N-terminal and C-terminal domains of HCS
2. Determine the effect of the MCD and novel mutants on the interaction determined in aim 1
3. Solve the structure of the proposed N-terminal structured domain using NMR
4. Characterise the interaction between wild type and mutant HCS and its substrate hPC107

CHAPTER 2

General Methods & Materials

2.1 Materials

2.1.1 General materials

Materials	Manufacturer
Amicon® Centrifugal Filter Devices (10,000 MWCO)	Millipore, MA, USA
Durapore Membrane Filter	Millipore, MA, USA
HisTrap	GE Healthcare, Buckinghamshire, England
Minisart syringe filter 0.45 µm or 0.80 µm	Sartorious, Goettingen, Germany
Nunc MaxiSorp™ flat-bottom 96 well plate	Nunc, Roskilde, Denmark
PVDF membrane (Hybond P)	GE Healthcare, Buckinghamshire, England
Glutathione-agarose	Scientifix, Australia

2.1.2 Chemical reagents

All chemicals and reagents were of analytical grade or of the highest purity available. Most common laboratory chemicals were purchased from Sigma-Aldrich Inc (St Louis, MO, USA) or BDH Chemicals Ltd. (Victoria, Australia). Specialised reagents and their suppliers are listed below.

Reagents	Supplier
1kb and 100bp DNA ladders	New England Biolabs, MA, USA
Bradford Protein Reagent Concentrate	Bio-Rad Laboratories Inc., CA, USA
d-Biotin	Sigma-Aldrich, MO, USA
d-[8,9-3H] Biotin	PerkinElmer, MA, USA
T4 DNA ligase	New England Biolabs, MA, USA
VENT DNA polymerase	New England Biolabs, MA, USA
Antarctic Phosphatase	New England Biolabs, MA, USA
Reduced Glutathione	Sigma-Aldrich, MO, USA
Optiphase Hi Safe Scintillation Fluid	Wallac (Perkin Elmer), MA, USA
Thrombin (From human plasma)	Sigma-Aldrich, MO, USA
Phusion DNA polymerase	New England Biolabs, MA, USA
PMSF	Amresco
Dithiothreitol (DTT)	Sigma-Aldrich, MO, USA
Tritiated Biotin	PerkinElmer, MA, USA
IPTG	BioVectra, PE, USA
MOPS / MES SDS running buffer	Invitrogen Life Technologies Inc, NY, USA
Restriction endonucleases	New England Biolabs, MA, USA
SeeBlue® Plus2 Prestained Protein Marker	Invitrogen Life Technologies Inc, NY, USA

2.1.3 Restriction endonucleases

All restriction endonucleases, provided with appropriate buffer concentrates, were purchased from New England Biolabs, MA, USA.

2.1.4 Antibodies

Antibodies were reconstituted and stored as per manufacturers instructions. Antibodies were used at the manufacturers recommended dilutions.

Reagents	Supplier
Cy2 conjugated AffiniPure Donkey Anti-goat IgG	Jackson ImmunoResearch Laboratories, PA, USA
Cy2 conjugated AffiniPure Donkey Anti-rabbit IgG	Jackson ImmunoResearch Laboratories, PA, USA
Cy5 conjugated AffiniPure Donkey Anti-rabbit IgG	Jackson ImmunoResearch Laboratories, PA, USA
Anti-GST antibody	Booker Laboratory, University of Adelaide

2.1.5 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, USA).

E. coli BL21: *E. coli* B F^- *dcm ompT hsdS*(r_B^- m_B^-) *gal* and *E. coli* Rosetta (DE3) for expression of recombinant proteins (Stratagene, La Jolla, CA, USA).

E. coli birA85 bioC strain BM4062 for complementation assay (Barker & Campbell, 1981).

E. coli Rosetta (DE3)pLysS^{RARE} (Novagen/EMD4Biosciences, USA) for expression of pET HCS proteins

2.1.6 Bacterial media

Luria Broth - 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% NaCl, adjusted to pH 7.0 with 5N NaOH.

LB agar - LB supplemented with 1.5% (w/v) bacto-agar.

2x YT - 1.6% (w/v) Bacto Tryptone, 1% Bacto yeast extract, 0.1M NaCl, adjusted to pH 7.0 with 5N NaOH.

Bacterial selection - Selection of bacteria bearing plasmid was achieved through the addition of appropriate antibiotics to both liquid and solid media. Ampicillin was used at 100 µg/mL and Chloramphenicol at 30 µg/mL.

2.1.7 Yeast Strain

Stock cultures of this strain and transformants were stored as glycerol stocks at -80°C.

EGY48:

The *Saccharomyces cerevisiae* EGY48[p8opLacZ] strain (Estojak *et al.*, 1995) was used in transformations and was a host for recombinant plasmids in yeast-two-hybrid assay controls and screens. The EGY48[p8opLacZ] reporter host strain carries a wildtype LEU2 gene under control of LexA operators and a p8opLacZ 10.3 kb reporter plasmid that encodes a LacZ gene under the control of LexA operators (CLONTECH Laboratories, 1996). EGY48: MATα, his3, trp1, ura3, $LexA_{op(x6)}$-Leu2.

2.1.8 Yeast media

Synthetic-dropout selection (or stock) media contains yeast-nitrogen-base (final 0.8% w/v), carbon source (glucose: final 2% w/v) and synthetic-dropout amino acids (final 1×). In synthetic-dropout induction (or screen) media, the carbon source is galactose (final 2% w/v) supplemented with raffinose (final 1% w/v). Solid media is prepared from liquid media supplemented with 2% w/v Bacto-Agar. During yeast two-hybrid screening the synthetic-dropout/galactose/raffinose induction plates are supplemented with 10× BU salts (final 1×) and XGAL (final 80mg/L). Note: If medium is too hot (>55°C) when BU salts added they will precipitate; XGAL is thermolabile and will be destroyed if added to hot medium.

2.1.8.1 Synthetic-Dropout Amino Acid Mix

Synthetic-dropout media is made with omission of appropriate amino acids in the 10× stocks that contain the following ingredients per 100 mL:

20 mg each of arginine, histidine (omitted for selection of plasmid pLexA) and methionine.

30 mg each of isoleucine, lysine and tyrosine.

50 mg each of phenylalanine, leucine (omitted for selection of protein-protein interaction in screen) and tryptophan (omitted for selection of plasmid pB42AD).

100 mg of uracil (omitted for selection of plasmid p8op-LacZ).

150 mg each of valine and adenine.

200 mg of threonine.

10× amino acid stock is prepared by dissolving all of the above ingredients in Milli-Q® water at 70-80°C, with stirring and sterilised by filtration through a Sartorius Minisart 0.45 µM filter. Stocks are stored at 4°C. Adenine and Uracil are added with the amino acids, instead of separately, for convenience.

2.1.9 Oligonucleotide Primers

All primers were purchased from Geneworks Pty Ltd., Hindmarsh, South Australia. All primers used were of sequencing grade.

Primer Name	Sequence 5' – 3'	Use
B92	tacatctaccccgggctcgagttaccgccg	Reverse HCS 1-726 with *XhoI* and stop codon
B106	gctctctcttagatttgtttcatcc	Internal forward HCS sequencing primer
B107	gagataatcggctcttaagg	Internal reverse HCS sequencing primer
B117	atctacgaattcatggaagatagactccacatgg	Forward Met1HCS with *EcoRI* site for pLexA
B118	atctacctcgagttacctgtccttgtcctcattctcc	Reverse HCS 1-314 with *XhoI* and stop codon
B212	atctacgaattcggaaaggcacccaacatcctcc	Forward G159 with *EcoRI*
B213	gtagatctcgagttacgtgaggttgactctcctcc	Reverse HCS 1-158 with *XhoI* and stop codon
B225	gatccccaggagatatcaagcttg	Annealed oligos to correct GST-NHCS fusion constructs
B226	gggtcctctatagttcgaacttat	Annealed oligos to correct GST-NHCS fusion constructs
B273	gtagatctcgagttaatgatgatgatgatgatgttcctgtgctttcaggttttcc	Introduction of TEV site in HCS
B277	atctacggatccggaaaggcacccaacatcctcctc	Forward primer for G159-E292 or G159-D349 HCS
B278	gtagatctcgagttattcctggtacctgcagccac	Reverse primer for G159-E292 NHCS
B279	gtagatctcgagttaatcttctggagtttgcsctatgttggagc	Reverse primer for G159-D349 NHCS
pLexA-5´	cgtcagcagagcttcaccattg	pLexA sequencing primer forward
pLexA-3´	cgcccggaattagcttggctg	pLexA sequencing primer reverse
pB42AD-5´	ccagcctcttgctgagtggagatg	pB42AD sequencing primer forward
pB42AD-3´	ctttctggcaaggtagacaagccg	pB42AD sequencing primer reverse
USP	cgccagggttttcccagtcacgac	Universal sequencing primer forward (pGEM-T plasmids)
RSP	cacacaggaaacagctatgaccatc	Universal sequencing primer reverse (pGEM-T plasmids)

2.1.10 Cloning vectors

2.1.10.1 Bacterial Cloning Vectors and Protein Expression Vectors

pGem-T EASY (Promega): 3015 bp cloning vector for cloning PCR products.

pGex4T2 (AP Biotech): 4900 bp cloning and bacterial expression vector; used to generate fusions of a known protein with glutathione-S-transferase (GST) protein (26 kDa).

pET32a^{+} (Novagen): 5900 bp cloning and bacterial expression vector; used to generate fusions of a known protein with Thioredoxin (Trx) (109aa) and 6-His tag.

2.1.10.2 Yeast Two-Hybrid Vectors (CLONTECH Laboratories, 1996)

pLexA: 10.2 kb cloning and yeast expression vector (Gyuris, 1993); used to generate fusions of the target protein with the LexA protein (amino acids 1-202).

pB42AD: 6.45 kb cloning and inducible yeast expression vector (Gyuris, 1993); used to generate fusions of a known protein (or a collection of library-encoded proteins) with the B42 AD and HA epitope-tag (YPYDVPDYAS).

p8op-LacZ: 10.3 kb reporter plasmid; encodes a LacZ gene under control of LexA operators (Estojak *et al.*, 1995).

pLexA-53: 11.1 kb positive control plasmid; encodes lexA/murine p53 (amino acids 72-390) fusion protein in pLexA.

pB42AD-T: 8.5 kb positive control plasmid; encodes an AD/SV40 large T-antigen (amino acids 87-708) fusion protein in pB42AD.

2.1.11 Commercial kits

Kit	*Supplier*
pGEM-T Easy Vector System I Kit	Promega, WI, USA
QIAprep Miniprep Kit	QIAGEN, GmbH, Germany
QIAquick Gel Extraction Kit	QIAGEN, GmbH, Germany
QIAquick PCR Purification Kit	QIAGEN, GmbH, Germany
QuikChange Site-Directed Mutagenesis Kit	Stratagene, CA, USA

2.1.12 Buffers and solutions

Blocking Solution (for Westerns): 1% BSA (w/v) in PBS containing 0.05% (v/v) Tween20

Coomassie Blue stain: 0.2% (w/v) Coomassie brilliant blue, 10% (v/v) Methanol , 10% (v/v) Acetic Acid.

Coomassie Destain: 10% (v/v) Methanol, 10% (v/v) Acetic Acid.

Immuno Blocking Solution: PBS containing 5% Horse serum.

Loading Buffer (for DNA) 6X: 0.5x TBE, 40% (v/v) glycerol, 1 mg/mL bromophenol blue.

PBS: 0.137 M NaCl, 2.7 mM KCl, 1.46 mM KH2PO4, 8.1 mM Na2HPO4 (pH 7.4).

PBS-Tween (PBS/T): PBS, 0.05% (v/v) Tween20.

SDS-PAGE Sample Buffer (for Proteins) 5X: 0.25 M Tris (pH 6.8), 10% (w/v) SDS, 0.5% (w/v) bromophenol blue, 50% (w/v) glycerol.

TAE: 40 mM Tris, 20 mM NaAc, 10 mM EDTA (pH 8.2)

TBS: 25 mM Tris pH 7.5, 150 mM NaCl.

TE: 10 mM Tris pH 7.5, 1 mM EDTA.

Towbin Transfer Buffer: 20 mM Tris-HCl pH 7.5, 1.15 M Glycine, 20% Methanol, 0.1% (w/v) SDS.

X-Gal: 20 mg/mL in DMF, stored at -20°C

2.1.13 Plasmids and vectors

Plasmids obtained for use in this study:

Yeast Two Hybrid

pLexA Matchmaker Yeast Two Hybrid System, Clontech, CA, USA.

pLexA-53 Matchmaker Yeast Two Hybrid System, Clontech, CA, USA.

pB42AD Matchmaker Yeast Two Hybrid System, Clontech, CA, USA.

pB42AD-T Matchmaker Yeast Two Hybrid System, Clontech, CA, USA.

pET (hBPL-His6) Dr. Steven Polyak, University of Adelaide, Australia.

pGEX(hPC107) Dr. Steven Polyak, University of Adelaide, Australia.

2.1.14 Computer software

Data were analysed using Microsoft Excel 2010 software and Graphpad Prism 5. ApE was used a plasmid map drawing software. UCSF Chimera was used for viewing and analysis of PDB files.

2.1.15 Web resources

NCBI (http://www.ncbi.nlm.nih.gov/) was used to access protein, nucleotide and PubMed databases and the BLAST (Basic Local Alignment Search Tool) databases. Multiple sequence alignments were performed using ClustalW (http://www.ebi.ac.uk/clustalw/). The ExPasy Proteomics Tools server (http://au.expasy.org/tools/) provided access to protein prediction programs (Phyre, Fugue and LOOPP).

2.2 General methods

2.2.1 Determination of protein concentration

Protein concentration was assayed using either the Bradford Reagent (Bio-Rad Laboratories Inc., CA, USA) or BCA Assay (Pierce, IL, USA) method depending on sample buffer compatibility. A standard curve of bovine serum albumin (BSA) was generated from 0 to 1 mg/mL. For the Bradford assay 10 µL of sample was mixed with 200 µL of 1x Bradford Reagent in a 96 well plate (Falcon). Absorbance at 620 nm wavelength was measured on a microplate reader (Molecular Devices, CA, USA). Standard curves were generated and linear regression used to calculate protein concentration using Microsoft Excel 2003.

2.2.2 Purification of GST-tagged proteins using Profinia™

Purification of GST-tagged protein was performed using the Profinia™ Protein Purification System (BIO-RAD) as specified by the user manual up to the wash step. The Bio-Scale™ Mini Profinity™ GST 5 ml cartridge was used. The purification was carried out as specified by the user manual. The buffer employed was TBS, pH 8.5. On column cleavage was performed by injecting 5 ml TBS, pH 8.5 containing 30 U of thrombin and 2.5 mM $CaCl_2$. Thrombin cleavage was performed at 37°C for 2 hours followed by the elution of cleaved protein using TBS, pH8.5.

2.2.3 SDS-PAGE electrophoresis and gel staining

Protein samples containing 5x Protein Sample Buffer were boiled for 5 minutes and then centrifuged briefly to collect condensation. Protein samples were loaded and fractionated on NuPage® 4-12% Bis-Tris polyacrylamide gels (Invitrogen) using 1X NuPAGE® MES Running Buffer (Invitrogen) at 200 V for approximately 40 minutes or until the dye front reached the bottom of the gel.

2.2.4 Western Blotting

Proteins were transferred onto a PVDF membrane using a semi-dry transfer unit (Hoefer SemiPhor, Amersham Pharmacia Biotech, CA, USA). Six sheets of Whatman filter paper and the PVDF membrane were pre-soaked in Towbin Transfer Buffer prior to assembly of the gel sandwich. Proteins were transferred for 1 hour at 80 mA per gel. The membranes were then blocked in 1% (w/v) BSA blocking buffer for 1 hour at room temperature (or at 4°C overnight) and washed 3 times with PBS-Tween. The membrane was probed with primary antibody (diluted in PBS-Tween) for 1 hour at room temperature and washed 3 times with PBS-Tween. The primary antibody used was rabbit anti LexA DNA binding domain (diluted 1:1000) or goat anti HA (diluted 1:1000), for either the LexA fusion proteins or the pB42AD fusion proteins, respectively. The membrane was probed with secondary antibody (diluted in PBS-Tween) for 1 hour at room temperature and washed 3 times with PBS-Tween. The secondary antibodies used were Cy5 anti rabbit (diluted 1:5000) and Cy2 anti goat (diluted 1:5000). The probing was performed concurrently by using either both primary or both secondary antibodies and using a single membrane for each experiment. The blot was visualised using the Typhoon TRIO Variable Mode Imager.

2.2.5 HCS activity assays

For *in vitro* HCS assays, the apo-biotin domain hPC107 (human pyruvate carboxylase 107) was used. *In vitro* biotinylation assays were carried out as described previously (Chapman-Smith *et al.*, 1999). Briefly, to quantitate HCS activity an *in vitro* biotinylation reaction was performed for one hour at 37°C in a 20 µL reaction mix containing 50 mM Tris pH 8.0, 3 mM ATP, 5.5 mM $MgCl_2$, 0.1 mg/mL BSA, 0.1 µM DTT, 5 µM apo-hPC107, 10 µM ^{3}H biotin and 100 µg lysate. At completion of the assay aliquots were spotted onto Whatman filter paper pre-treated with 10% TCA and 0.1 mg/mL biotin. Filter papers were washed for 20 min in ice-cold 10% TCA and once in 10 min 100% ice-cold ethanol, dried,

and added to 2 mL Optiphase HiSafe Scintillation fluid. Radioactivity measurements were obtained using a Rackbeta Liquid Scintillation Counter (Perkin Elmer, MA, USA).

2.2.6 Circular Dichroism

CD was performed using a Jasco J-185 CD spectrophotometer. Individual spectra (including 10 mM sodium phosphate buffer, pH 8.0, as a blank) were recorded at 20°C across wavelengths 185 nm to 300 nm with a nitrogen flow of 7.5 L/min, scan speed of 20 nm/min and accumulation mode of 5 (average of 5 scans).

2.2.7 Molecular biology techniques

2.2.7.1 Primer Design

Primers for polymerase chain reaction (PCR) reactions were designed with the aid of Oligo™ software. Generally, primers were >17 bp long, terminated in a G/C residue, had >40% G/C content, unique 3' end of 7 nucleotides (compared to the template) and did not self anneal. Primer pairs had similar melting temperature at which half maximal binding occurs. Primers designed to amplify sequences for cloning incorporated a restriction site at the beginning of the coding sequence (generally *Eco*RI) and a different restriction site (generally *Xho*I) and stop codon at the end of the coding sequence. Primers designed for mutagenesis reactions incorporated nucleotide mismatch(es) in the centre flanked on both sides by at least 10 matched nucleotides. Where possible, mutagenesis primers introduced or removed a restriction site for later ease of screening for mutants.

2.2.7.2 PCR protocols

PCR reactions routinely consisted of the following in a 50 µL reaction: 1x ThermoPol Buffer, 250 ng oligonucleotide primers, 400 µM dNTP mixture and 1 U VENT DNA polymerase. Reactions were routinely carried out under the following cycling conditions: 30

seconds denaturation at 95°C, 30 seconds annealing at 60°C, 90 seconds extension at 72°C for 25-35 cycles. All reactions were carried out on a MJ Research PTC 2000 Thermal Cycler (GMI Inc., MN, USA).

2.2.7.3 Site-Directed Mutagenesis

The Quikchange™ mutagenesis kit (Stratagene) was used for mutagenesis PCR reactions according to the manufacturer's recommendation. *Pfu* DNA polymerase was used to replicate both plasmid strands with high fidelity and without displacing the mutant oligonucleotide primers, which are each complementary to opposite strands of the vector (Stratagene). Briefly, each 50 µL reaction contained 1× reaction buffer, 1 mM dNTPs, 2.5 ng/µL each primer, 2.5 U *Pfu* DNA polymerase and 0-50 ng of template DNA. Negative control reaction contained no *Pfu* DNA polymerase. Initial denaturation (94°C) for 5 min was followed by 16-20 cycles of denaturation (94°C) for 1 min; annealing (50-60°C) for 1 min; and extension (68°C) for 9 min, in a PTC-100 Programmable Thermal Cycler (MJ Research Inc). Samples were stored at 4°C until *Dpn*I digestion, which removed parental DNA. Reactions were precipitated in NaAc/EtOH, resuspended in Tris-EDTA (TE) and transformed in to DH5α competent cells using heat shock or electroporation. Mutants were identified, where possible, using restriction digest when mutagenesis primers introduced or removed a restriction site; otherwise, they were identified by DNA sequence analysis. The DNA sequence of the mutant clones was verified by automated DNA sequence analysis.

2.2.7.4 Phosphorylation and annealing of Oligonucleotides

Oligonucleotides were dissolved in sterile water at a dilution of 1 µg/µl. This was then diluted 1 part in 10 to give a working stock of 100 ng/µl. Oligonucleotide (100 ng) was incubated with 5 mM ATP and 1U T4 polynucleotide kinase in 1X PNK buffer (Geneworks, SA, Australia) for 1 hour at 37°C, followed by inactivation of the enzyme at 65°C for 5 mins.

Complementary oligonucleotides (50 ng of each) were subsequently combined, boiled for 5 mins then left to cool to room temperature for 20 mins.

2.2.7.5 Agarose gel electrophoresis

Analysis of DNA and separation of DNA fragments was performed using agarose gel electrophoresis. Gel slabs were poured by melting 0.7-2% (w/v) agarose in TBE buffer. Prior to loading into wells DNA samples were mixed with an appropriate volume of 6X DNA Loading Buffer. Samples were electrophoresed in TBE buffer at 100-150V and then stained in Ethidium bromide solution (1 µg/mL) for 10 minutes followed by destaining in distilled water. DNA was visualised on a UV transilluminator and photographed using a Mitsubishi Video Processor.

2.2.7.6 Preparation of DH5α competent cells

A single colony was used to inoculate a 2 mL overnight culture of DH5α in Luria Broth. After incubation overnight at 37°C with rotation, 330 µL was subcultured into 10 mL of LB and the culture grown until it reached log phase (as indicated by an OD600nm of 0.5-0.6). Five mL of log phase culture was then added to 100 mL pre-warmed LB and incubated for a further 1.5 hours. The culture was split into 4 x 50 mL centrifuge tubes and placed on ice for 5 minutes. Cells were pelleted by centrifugation at 3,200 g for 5 minutes at 4°C. The cells were then resuspended in a total of 40 mL of Transformation Buffer 1 and incubated on ice for 5 minutes, then collected by centrifugation at 3,200 g for 5 minutes at 4°C. This was followed by resuspension in Transformation Buffer 2 to a final volume of 4 mL. Cells were chilled on ice for 5 minutes prior to aliquoting and storing at -80°C until required.

2.2.7.7 Restriction digest of DNA

Routinely, 1-5 µg of DNA was digested with 1-10 U of restriction enzyme in the appropriate NEB buffer for 2 hours at 37°C. In the instance of double digests the buffer conditions for the more sensitive enzyme were used (as recommended by NEB). For cloning, DNA fragments were separated by agarose gel electrophoresis before purification from the excised gel slice using either a QIAGEN.

2.2.7.8 Ligation of DNA fragments

Ligation was carried out in a 10 µL volume in an insert:vector ratio of 3:1 in 1X ligase buffer and 2 U of T4 DNA ligase for 1 hour at room temperature.

2.2.7.9 Transformation of competent cells

Five µL of ligation mixture was added to 100 µL of competent cells and incubated on ice for 30 minutes. Cells were incubated at 42°C for 5 minutes then allowed to recover for 5 minutes on ice. Cells were immediately plated onto pre-warmed LB agar plates for ampicillin selection, otherwise a 1 hr recovery step with 1 mL LB was performed prior to plating.

2.2.7.10 Glycerol stocks

For long term storage of *E. coli* strains, equal volumes of an overnight culture and 80% glycerol were mixed and stored at -80°C.

2.2.7.11 Purification of plasmid DNA

For purification of small (<10 µg) amounts of plasmid DNA the QIAGEN QIAprep Miniprep Kit was employed. For larger scale purification the QIAGEN Midi prep Kit was used according to manufacturer's instructions. For transfection quality DNA, purified material

from the kits was precipitated in 70% (v/v) isopropanol, pelleted by centrifugation at 17,500 g for 30 minutes and then washed in 70% (v/v) ethanol before resuspension in sterile MilliQ water.

2.2.7.12 Quantification of DNA

DNA was quantified by measuring absorbance at 260 nm using a CARY WinUV Spectrophotometer. A measure of 1 OD260 unit is equal to 50 µg/mL.

2.2.7.13 DNA sequencing

Plasmid DNA or PCR products were used as templates for DNA sequencing. A 20 µL reaction containing 4 µL DNA, 1 µL 100 ng of appropriate primer, and 1 µL BigDye (version 3) reaction mix (Perkin Elmer, Applied Biosystems, CA, USA), 4 µL 5x buffer and 10 µL autoclaved water was prepared. The PCR profile consisted of 30 cycles of denaturation at 96°C for 30 seconds, annealing at 50°C for 15 seconds and extension at 60°C for 4 minutes. After thermocycling, 80 µL of 75% (v/v) isopropanol and 1 µL glycogen (20 mg/mL) were added to the PCR products, vortexed and left at room temperature for 30 minutes. Precipitated DNA was isolated by centrifugation at 17,500 g for 20 minutes. The pellet was washed twice in 75% isopropanol followed by centrifugation for 5 minutes at 17,500 g. The pellet was then dried in a 37°C heating block prior to being submitted to the Molecular Pathology Sequencing Service at the Institute of Medical and Veterinary Science, Adelaide. Sequencing was carried out on a 3700 Applied Biosystems Analyser.

2.2.8 Yeast methods

2.2.8.1 Preparation of competent *S. cereviseae*

A 50 mL overnight culture of yeast in appropriate dropout media was prepared (depending upon the auxotrophic marker of plasmid) by inoculating the media with the colony of interest and incubating the culture overnight at 30°C with shaking. The cells were harvested at 4200 x g for 5 min using a Beckman centrifuge and the pellet was resuspended in 50 mL YP. The culture was incubated at 30°C with shaking for a further 4 hours. The cells were harvested at 4200 x g for 5 minutes and washed in 25 mL sterile MQ, centrifuged at 4200 x g for 5 min. The pellet was resuspended in 1 mL 100 mM LiAc, transferred to a 1.5 mL eppendorf tube and centrifuged at 17 900 x g for 15 sec using a bench top microcentrifuge. The LiAc was removed and the cells were resuspend in 200 µL 100 mM LiAc. The competent cells were used immediately as loss of competency occurs after a few hours.

2.2.8.2 Transformation of competent yeast strains

The competent cells were vortexed and 50 µL of cells were transferred to fresh Eppendorf tubes. To the cells in a total volume of 350 µL the final concentration of reagents that were added were 35 % (w/v) PEG and 0.145 M LiAc. Five µL of herring sperm DNA (10 mg/mL) and 10 µL of the appropriate plasmid (5-10 µg) were added. The tubes were vortexed vigorously then incubated with shaking at 30 °C for 30 minutes. The cells were subsequently heat shocked at 42 °C in a water bath for 25 minutes. The cells were centrifuged at 2800 x g for 3 minutes using a bench top microcentrifuge. The supernatant was removed and the cells were resuspended in 200 µL sterile MQ, and 100 µL of the cells were spread plated onto the appropriate selection media. The plates were incubated 30 °C for 3-4 days.

CHAPTER 3

Exploring the inter-domain interactions within HCS

3.1 Introduction

3.1.1 Investigating the inter-domain interactions within HCS

The first hypothesis of this study is that the N-terminal half of HCS interacts (N-HCS) with the C-terminal region of HCS (C-HCS). Additionally, I will test whether the proposed structured N-terminal domain (159-314 N-HCS) alone interacts with the C-terminal region of HCS. Interestingly, the V_{MAX} mutants that give rise to MCD all map within the domain G159-R314 N-HCS. I will then test two separate aims involving these mutants. First, whether the mutants disrupt an interaction between the two halves of HCS. Secondly, whether the mutants interfere with binding of the substrate biotin domain. These two protein:protein interactions can be investigated genetically using a Yeast Two-hybrid assay.

The MATCHMAKER LexA Two-Hybrid System (CLONTECH Laboratories, 1996) was used here to investigate protein:protein interactions. In this system the interaction between a target or "bait" protein (fused to the DNA binding domain or DNA-BD) and a "prey" or library-encoded protein (fused to the activation domain or AD) results in a novel transcriptional activator being produced that has binding affinity for LexA operators. This factor proceeds to activate reporter genes having upstream LexA operators. If no interaction occurs, then the reporter genes will not be transcribed. This Lex-A based two-hybrid system uses growth selection, the conditional expression of a nutritional reporter gene in order to screen large numbers of yeast transformed with the constructed fusion library for interacting proteins. In the MATCHMAKER LexA Two-Hybrid System, two expression vectors are used, namely pLexA and pB42AD. These were used to generate fusions of the DNA-binding domain (DNA-BD) and activation domain (AD), respectively. The MATCHMAKER LexA Two-Hybrid System was employed in this study due to the utilisation of dual selection. The system uses growth selection (expression of a nutritional reporter gene LEU2) and the transcription of GAL1. The GAL1 promoter in pB42AD (the

AD is an 88-residue acidic *E. coli* peptide (B42) that activates transcription in yeast) is used for the expression of cloned proteins fused to the B42AD. Transformants that contain the AD fusion plasmids are then grown in galactose containing medium, this is to induce expression of the fusion protein (colonies will be blue in media containing X-gal). Using this system was advantageous as it minimises false positives that can be found in systems that only utilise growth selection. The MATCHMAKER LexA Two-Hybrid System uses two reporter genes (LEU2 and lacZ) under the control of multiple LexA operator sites. An integrated LEU2 nutritional reporter gene results in the Leu– auxotrophic host yeast cell EGY48, to grow on induction medium that does not contain leucine when it is transformed with plasmids encoding interacting hybrid proteins. When lacZ transcription is activated in *S. cerevisea* strain EGY48 harbouring the plasmid reporter p8op-lacZ, the cells produce β-galactosidase.

3.2 Specific Methods

3.2.1 Nucleic acid manipulations

pLexA N-HCS or C-HCS and pB42AD N-HCS or C-HCS were cloned previously (by Dr Lisa Bailey, University of Adelaide). The oligonucleotides used to amplify DNA encoding M1-E151 N-HCS using PCR were B117 and B213 and the oligonucleotides used to amplify the proposed N-terminal structured domain G159-R314 N-HCS were B277 and B118 and pET HCS-His6 was the template DNA used. To amplify DNA encoding N-terminal domain mutant HCS the oligonucleotides employed were B117 and B118 and B117 and B92 for full-length HCS. L216R and L237P HCS mutants were PCR amplified from HCS that contained either pET-L216R HCS or pET-L237 HCS. The HCS mutants derived by error prone PCR, namely L166R, L206P, W210R, L246M, L270S, H306R, F321S and E181G/E327G, were PCR amplified using the B117 and B92 primers from their pAra templates. Similarly, the DNA encoding the N-HCS for this mutant series were amplified using oligonucleotides B117 and B118. The oligonucleotides employed in each PCR reaction were designed to engineer an *EcoR*I site onto the 5´ end or a *XhoI* site and a stop codon onto the 3´ end of the DNA. Purified PCR products were digested with *EcoR*I and *XhoI* and ligated into similarly treated pB42AD plasmid. All the ligations were transformed into (competent) *E. coli* DH5α *cells* and the plasmids were screened by restriction digest with *EcoR*I and *Xho*I. Clones producing the desired restriction pattern were confirmed by DNA sequencing using the pB42AD forward and reverse primers. The wild type HCS in pLexA and hPC107 in pB42AD were cloned previously (by Dr Lisa Bailey).

3.2.2 *In vivo* selection of HCS mutants

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 transcriptional regulator protein, *AraC*, is also contained in this vector. Random

mutagenesis was performed using error prone PCR using low fidelity *Taq* polymerase and 30 cycles of PCR. The error prone PCR was carried out on the DNA encoding the N-terminal region of HCS. The PCR mutated N-terminal HCS DNA mix was digested with *Eag*I and *Xba*I then ligated into similarly treated pAra(HCS-His6) vector. Error prone PCR and the ligation were performed by Dr Steven Polyak (University of Adelaide).

To select for amino acid substitutions that abolish HCS activity, the library of mutants was transformed into *E. coli birA*85*,* biotin auxotroph strain BM4062 (Barker & Campbell, 1981). The *birA85* mutation in BM4062 confers a temperature-sensitive phenotype to BPL. Under selective conditions, 42°C, only the cells expressing functional exogenous BPL survive, as *E. coli* acetyl CoA carboxylase can be activated. One thousand colonies were screened by picking the colonies from the master plate and streaked onto two LB (ampicillin) plates (100 colonies per plate). One plate was incubated at the permissive 30°C and the other at the restrictive 42°C. Clones that failed to grow at 42°C were rescued from the 30°C plate and analysed. Following re-testing, plasmids were extracted from the colonies that failed to complement using the QIAGEN® mini prep kit. The plasmids were sequenced.

3.2.3 Yeast Methods

Yeast culturing conditions, the media used, transformation methods and screening of transformants were performed as outlined in the Clontech Yeast Protocols Handbook for the MATCHMAKER LexA Two-Hybrid System with some modifications in the transformation protocol. In the manual, it recommends co-transformation of the pLexA or pB42AD fusion constructs. However, due to the low efficiency of co-transformation experienced, in this thesis sequential transformation was used where one plasmid was transformed at a time. All of the methods employed in the Yeast Two-hybrid for the preparation and transformation of competent yeast cells are outlined in Chapter 2, section 2.2.8.1-2.2.8.2.

3.3 Results

3.3.1 Determination of an interaction between the N-terminal and C-terminal regions of HCS using the Yeast Two-hybrid assay

As direct evidence supporting an interaction between N-terminal and C-terminal halves of HCS has not been reported, a Yeast Two-hybrid approach was initially performed to test this hypothesis. Here it was important to determine the optimal combination of the appropriate fusion proteins. In this study, the appropriate plasmids required were transformed into the *S. cerevisea* strain EGY48 harbouring the plasmid reporter p8op-lacZ. Three independent isolates of each strain were selected for the analysis. The cells were streaked onto an induction plate containing X-gal and BU salts (excluding uracil, histidine, trypophan and leucine). Blue staining colonies indicated a positive protein:protein interaction due to the reconstitution of a functional transcriptional activator required for expression of both β-galactosidase and leucine nutritional markers. No growth implied no protein:protein interaction. White colonies were false positives due to non-specific activation of the LEU2 reporter. These false positives contain AD proteins that non-specifically activate LEU2 gene whether or not the specific DNA-BD/target protein is present (Bartel *et al.*, 1993). This highlights the advantage of using a dual selection system over other Yeast Two-hybrid methods.

The positive control provided by the MATCHMAKER LexA Two-Hybrid System detected the interaction between p53 and SV40 large T-antigen (LexA p53 and B42AD T). Figure 3.1 shows that the positive control was able to grow on selection media and resulted in blue colonies in the presence of X-gal. In addition, the interaction between HCS fused to LexA and hPC107 fused to B42AD was tested and produced blue staining colonies on the selection media. The negative control of LexA (empty) and B42AD (empty) also resulted in white colonies but no blue staining. The combination that confirmed an interaction between the two regions of HCS was LexA (C-HCS) and B42AD (N-HCS). The LexA (N-HCS) and

B42AD (C-HCS) combination was negative as no growth and blue colour was observed. The negative result was most likely due to no expression or instability of the fusion proteins. For the rest of this study the combination that was used was LexA (C-HCS) and B42AD (N-HCS). In order to determine that C-HCS was not self-activating, a negative control containing LexA C-HCS and B42AD (empty) was used. An absence of growth demonstrated that LexA C-HCS was indeed not self-activating.

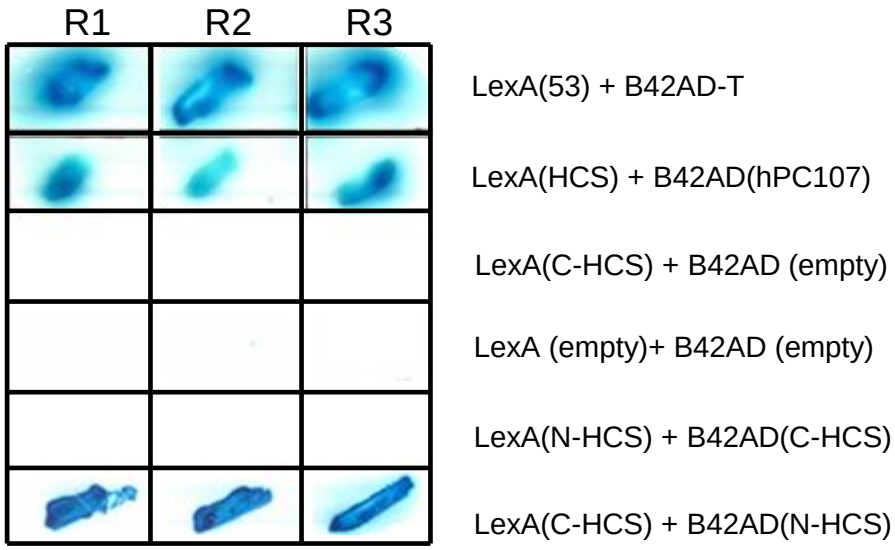


Figure 3.1 Determination of the optimal combination of the interaction between N-HCS and C-HCS using the Yeast Two-hybrid assay. R1-R3 grid shows three independent clones plated on induction/selection media (galactose, -URA, -HIS, -TRP, -LEU, +X-gal). Growth and blue colour indicates dual activation of LEU2 and LacZ reporter genes.

3.3.2 Constructs prepared for the interaction studies using Yeast Two-hybrid assay

After the optimal combination of the fusion proteins was determined the constructs that were required to define the region of N-HCS that interacts to C-HCS were prepared (figure 3.2(1A-B)). Figure 3.2(2A-B) shows the constructs that were made for the experiment that tested whether HCS mutants disrupt the interaction between N-HCS and C-HCS. Figure 3.2(3A-B) were the constructs made to determine the interaction between wild type and mutant HCS and hPC107. DNA sequencing was used to confirm each construct.

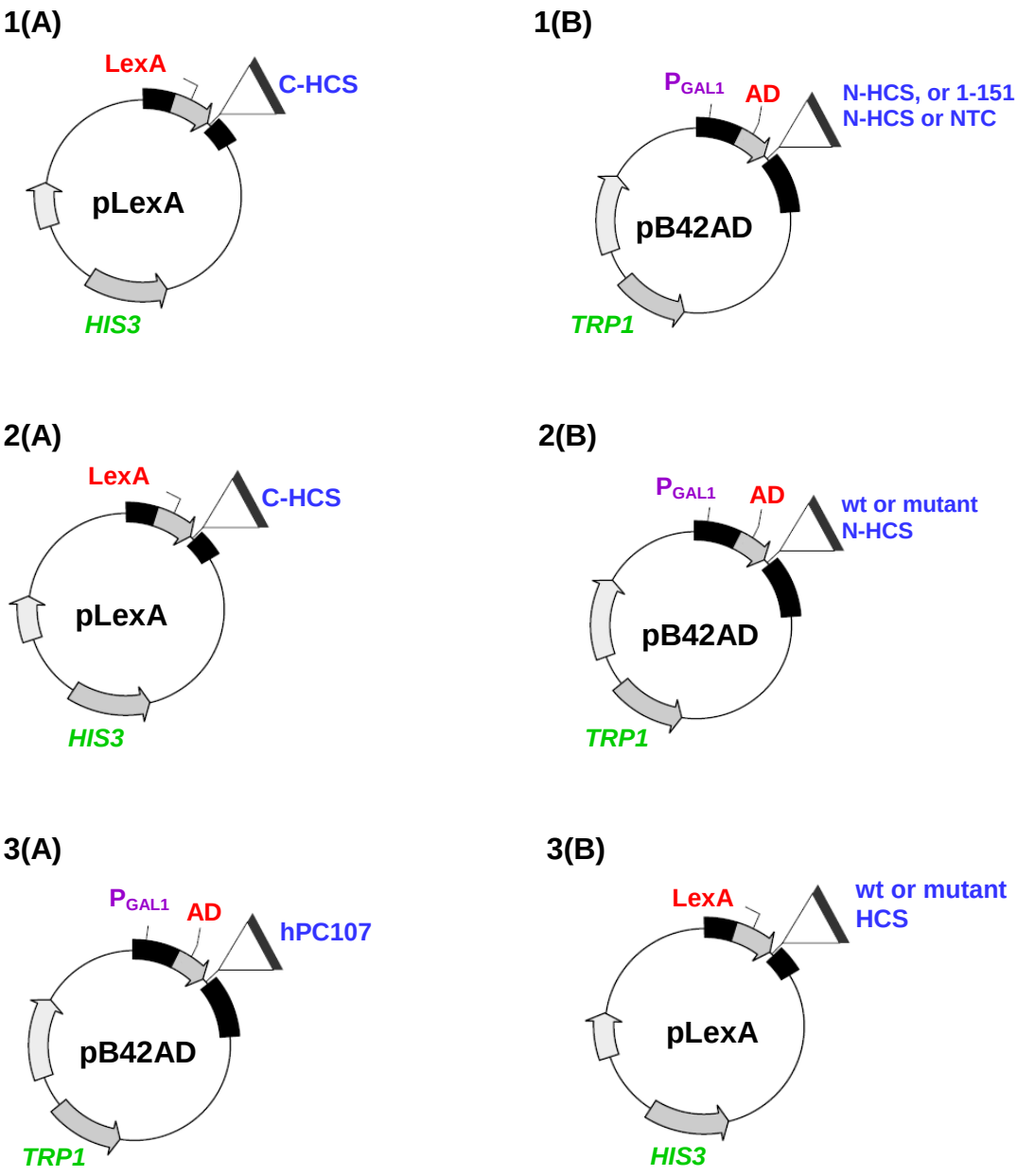
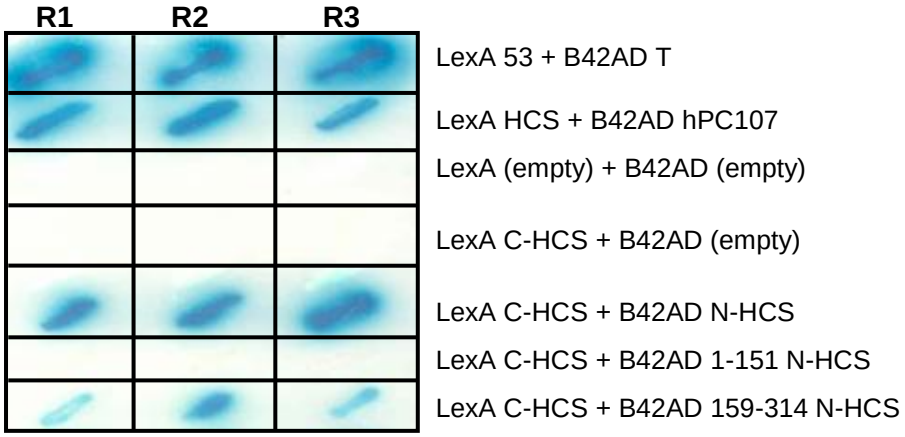


Figure 3.2 The constructs that were made for determining the interaction of **1(A)** C-HCS and **1(B)** N-HCS, 1-151 N-HCS or NTC; **2(A)** C-HCS and **(2B)** wild type or mutant N-HCS; **3(A)** hPC107 with **3(B)** wt and mutant HCS with hPC107 using Yeast Two-hybrid. Each insert was cloned using *EcoRI* and *XhoI* restriction sites.

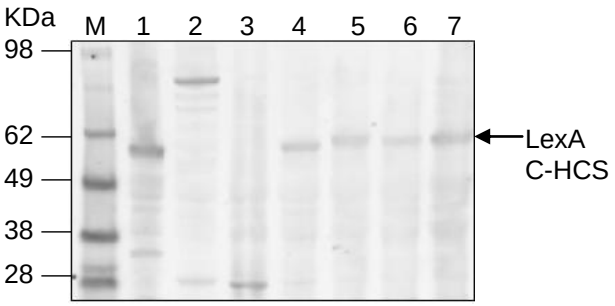
3.3.3 Defining the region of N-HCS that interacts with C-HCS

The appropriate constructs were transformed into EGY48[p8op-lacZ] yeast cells for the assay. Figure 3.3 (A) shows that while co-expression of LexA C-HCS and B42AD N-HCS resulted in growth and blue colouring on induction plates containing X-gal and BU salts, the construct encoding B42AD 1-151 N-HCS was unable to support growth. In contrast co-expression of the region encoding LexA 159-314 N-HCS and B42AD C-HCS was able to support growth and stained blue. These results indicated that 159-314 N-HCS interacted with C-HCS. Western blot analysis (figure 3.3(B)) performed on whole cell lysates failed to detect a product corresponding to residues 1-151 N-HCS, indicating that this fusion protein was either not expressed or was unstable in the yeast. All other fusions were detected in the Western blot analysis.

A



B(1)



B(2)

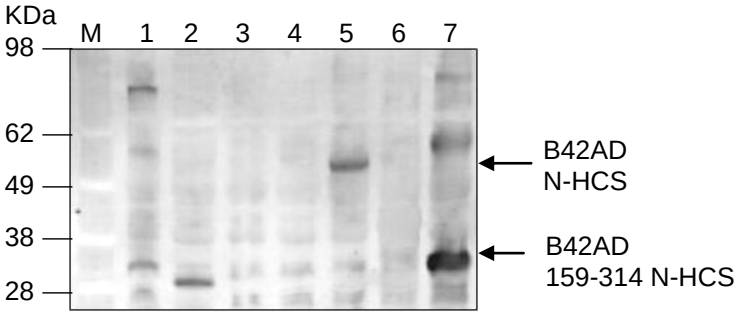


Figure 3.3 **(A)** Domain mapping of the interaction of N-HCS and C-HCS using the Yeast Two-hybrid assay. R1-R3 grid shows three independent clones plated on induction/selection media (galactose, -URA, -HIS, -TRP, -LEU, +X-gal). Growth and blue colour indicates dual activation of LEU2 and LacZ reporter genes.
(B) Western blot of Yeast Two-hybrid assay of interaction of wt and mutant N-HCS and CHCS. **B(1)** was probed with rabbit anti-LexA primary antibody and Cy5 anti-rabbit secondary antibody. **B(2)** was probed with goat anti-HA primary antibody and Cy2 anti-goat secondary antibody. A single membrane was used.
M - Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen)
1 - LexA 53 + B42AD T
2 - LexA HCS + B42AD hPC107
3 - LexA (empty) + B42AD (empty)
4 - LexA C-HCS + B42AD (empty)
5 – LexA C-HCS + B42AD N-HCS
6 - LexA C-HCS + B42AD 1-151 N-HCS
7 - LexA C-HCS + B42AD 159-314 N-HCS

3.3.4 *In vivo* complementation assay

Due to having only two V_{MAX} MCD mutants, namely L216R and L237P HCS, a larger pool of N-terminal mutations was generated to gain more information for the interaction studies. In order to generate a mutant library, error prone PCR mutagenesis targeted the proposed structured N-terminal domain. These novel mutant HCS were isolated using an *in vivo* complementation assay, which screened for mutants that caused HCS inactivation. The *E. coli* strain BM4062 contains the *birA85* mutation, which confers a conditionally lethal temperature-sensitive phenotype. At 42°C, the endogenous bacterial BPL is inactive, resulting in bacteria unable to grow. However, the growth defect can be overcome by recombinant expression of a functional HCS as BPLs from different organisms can biotinylate the bacterial apocarboxylase (Chapman-Smith and Cronan, 1999). The mutant library was transformed into BM4062 and grown at the non-restrictive temperature. Isolates were streaked onto duplicate plates and grown at either 30°C or 42°C. When no growth was observed at the restrictive temperature, the mutant HCS were deemed to be inactive. The positive control for growth was wild type HCS and the negative control was the parent vector pAra13. L237P HCS variant was used as a test control. Out of the one thousand screened colonies, 12 failed to complement the genetic defect. These were rescued off the duplicate plate grown at 30°C allowing extraction of plasmids for analysis. DNA sequencing of the isolated constructs showed that 4 isolates contained non-sense mutations, 7 produced a single amino acid substitution and 1 contained a double mutation. The N-terminal HCS mutants that were identified by DNA sequencing were L166R, L206P, W210R, L246M, L270S, H306R, F321S and the double mutant E181G, E327G HCS. Each mutant was then retested in the complementation assay. Figure 3.4 (A) shows that all of the isolates grew at the permissive temperature of 30°C, since the endogenous BirA was active at this temperature. At the restrictive temperature of 42°C (figure 3.4 (B)), only the positive control expressing wild type HCS was able to complement. As expected cells

harboring the parent vector pAra13 failed to complement. Similarly all of the isolated mutant HCS did not grow at the restrictive temperature. A sequence alignment against other eukaryotic species revealed that some of the mutations that affected HCS activity are conserved between all of the aligned sequences, suggesting their functional importance (figure 3.5).

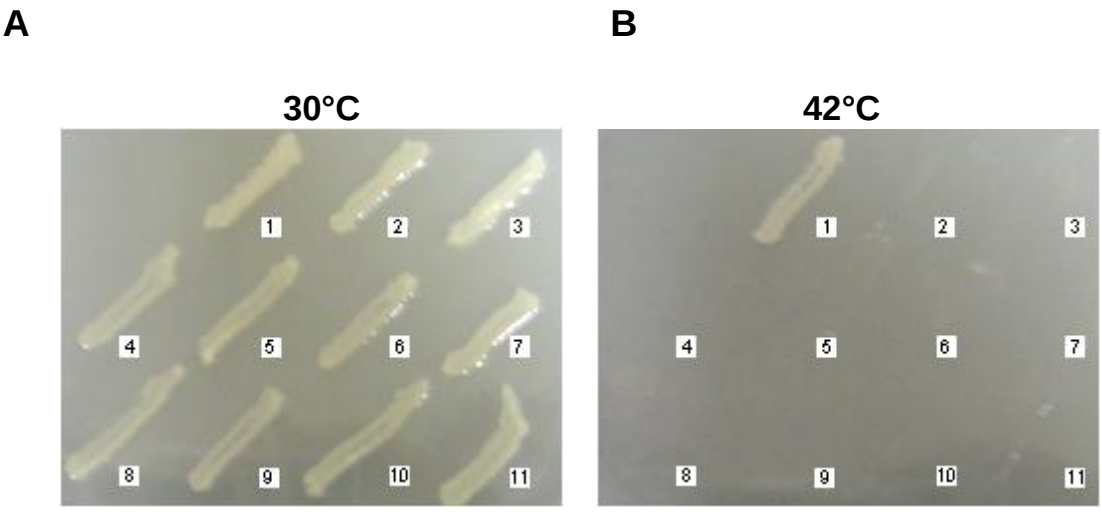


Figure 3.4 *In vivo* complementation assay for the isolation of non-functional N-terminal HCS mutants.
(A) Incubation at the permissive temperature of 30°C.
(B) Incubation at the restrictive temperature of 42°C.

1 - positive control wild type HCS
2 - negative control pAra13
3 - L237P HCS
4 - L166R HCS
5 - L206P HCS
6 - W210R HCS
7 - L246M HCS
8 - L270S HCS
9 - H306R HCS
10 - F321S
11 – double mutant E181G, E327G HCS

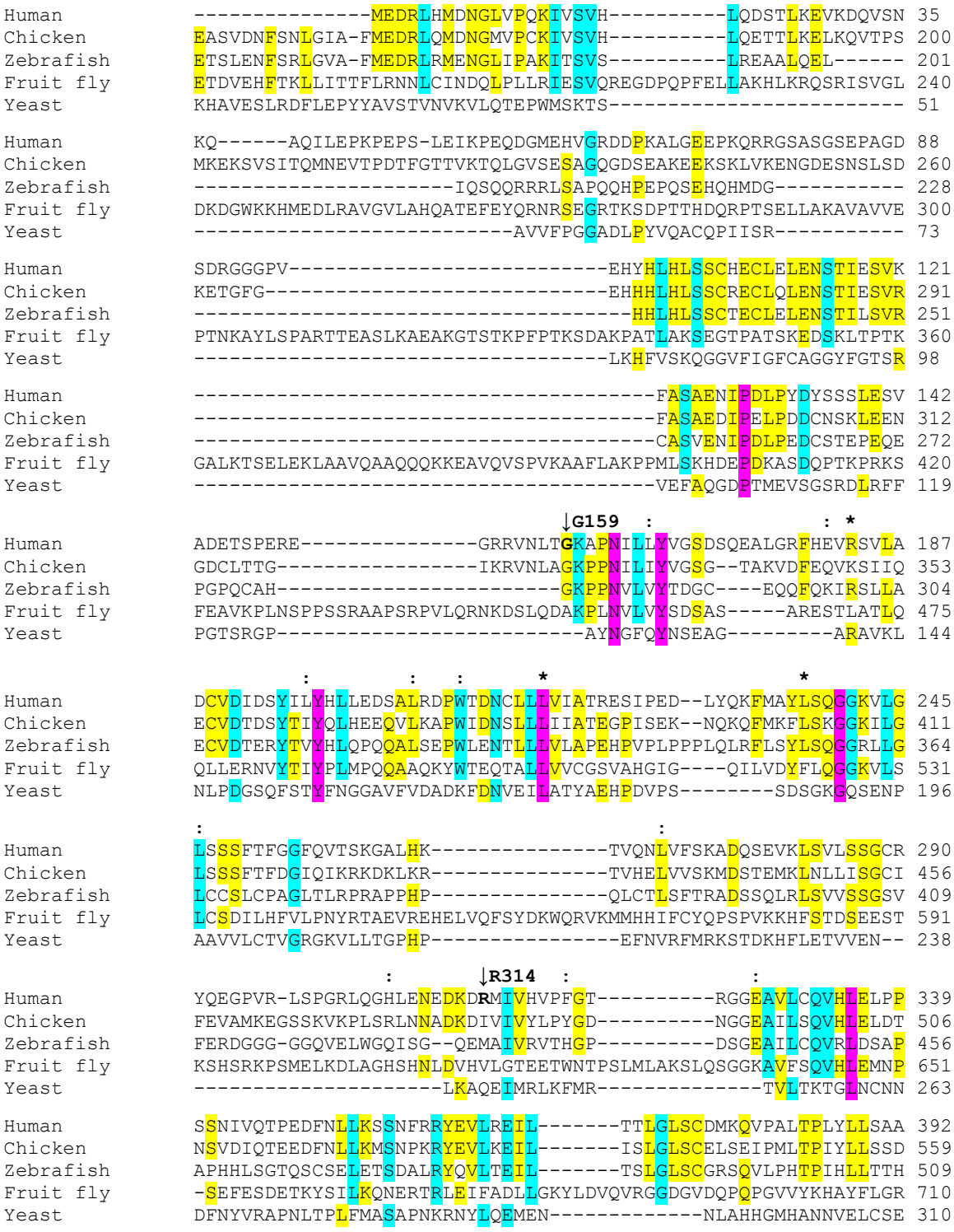


Figure 3.5 Alignment of the N-terminal region of eukaryotic holocarboxylases. Sequences from human, chicken, zebrafish, fruit fly and yeast were aligned in ClustalW. Residues highlighted with ***** show position of N-terminal HCS mutations that give rise to MCD and those highlighted with **:** show the novel mutants. The proposed N-terminal structured domain of HCS, G159-R314, is displayed in bold. Colours indicate the conservation of amino acids; ■ 100 %, ■ 80 %, and ■ 60 % identity.

3.3.5 Determining the effect of MCD and novel HCS mutants on the interaction between N-HCS and C-HCS domains

In section 3.3.3, the interaction between C-HCS and N-HCS was confirmed using the Yeast Two-hybrid assay. This interaction was further probed using the MCD and novel HCS mutants (figure 3.6). Here the positive controls were LexA HCS + B42AD hPC107 and LexA C-HCS + B42AD N-HCS. The negative controls were LexA (empty) + B42AD (empty) and the self activation control LexA C-HCS + B42AD (empty). With the negative control LexA (empty) and B42AD (empty) growth was observed, however blue staining was not observed confirming a lack of interaction. This highlights the importance of using a dual selection system in order to over come false positives. In all but one case the interaction between N-HCS and C-HCS was not affected by the MCD mutants nor the novel HCS mutants. The only exception was the H306R N-HCS mutant. The Western blot analysis demonstrated that all of the mutant N-HCS fusion proteins were expressed except H306R N-HCS (figure 3.7 (B). Therefore the apparent lack of an interaction between H306R N-HCS and C-HCS (figure 3.6, lane 10) was most likely due to a lack of expression of the fusion protein

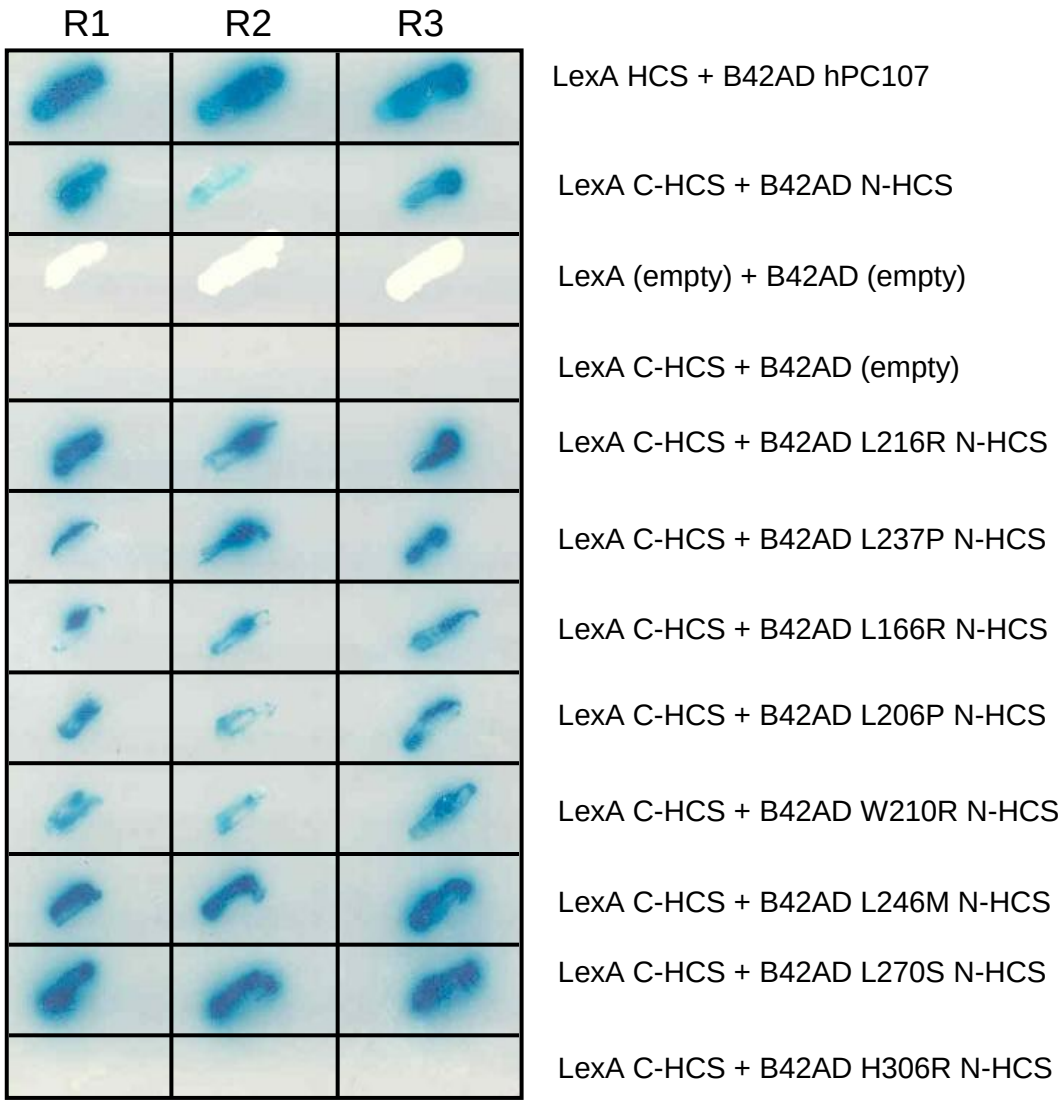


Figure 3.6 Effect of MCD and randomly generated novel mutations on the interaction of N-terminal and C-terminal HCS using Yeast Two-hybrid assay. R1-R3 grid shows three independent clones plated on induction/selection media (galactose, -URA, -HIS, -TRP, -LEU, +X-gal). Growth and blue colour indicates dual activation of LEU2 and LacZ reporter genes.

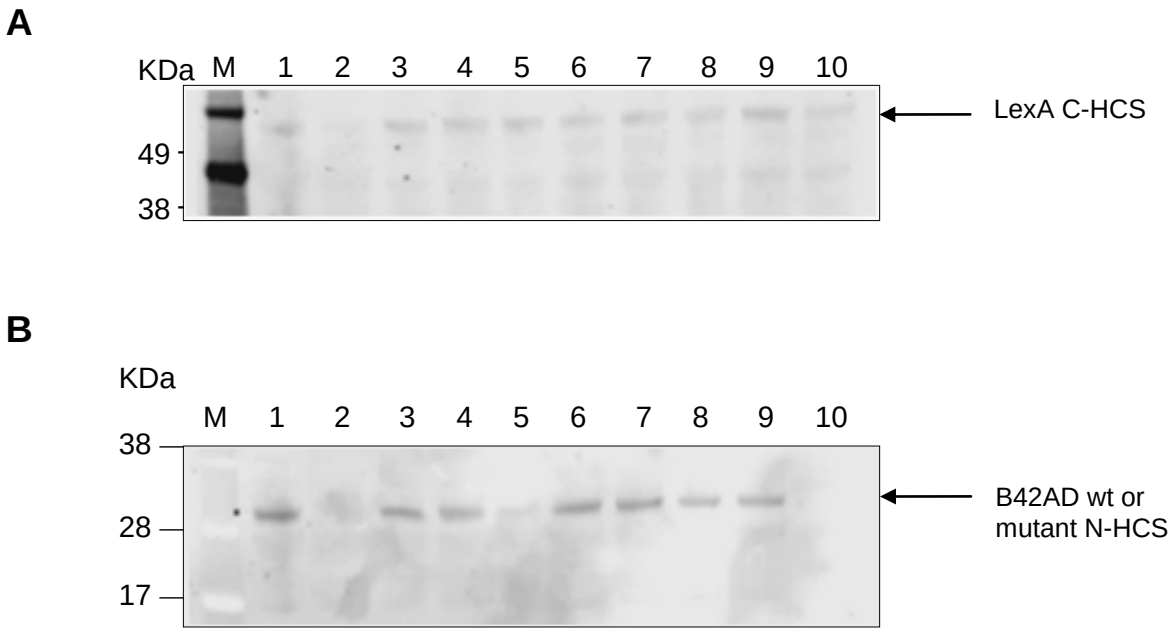


Figure 3.7 Western blot analysis to confirm expression of fusion proteins used in the Yeast Two-hybrid assay of interaction of wild type and mutant N-HCS with C-HCS.
(A) was probed with rabbit anti-LexA primary antibody and Cy5 anti-rabbit secondary antibody.
(B) was probed with goat anti-HA primary antibody and Cy2 anti-goat secondary antibody. A single membrane was used.
M is Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen)

1 - LexA CHCS + B42AD NHCS	6 - LexA CHCS + B42AD NHCS [L206P]
2 - LexA (empty) + B42AD (empty)	7 - LexA CHCS + B42AD NHCS [W210R]
3 - LexA CHCS + B42AD NHCS [L216R]	8 - LexA CHCS + B42AD NHCS [L246M]
4 - LexA CHCS + B42AD NHCS [L237P]	9 - LexA CHCS + B42AD NHCS [L270S]
5 - LexA CHCS + B42AD NHCS [L166R]	10 - LexA CHCS + B42AD NHCS [H306R]

3.3.6 Determining the effect of MCD and novel HCS mutants on the intermolecular interaction of HCS and hPC107

In section 3.3.5 I demonstrated that the MCD and novel HCS mutants did not disrupt the interaction between N-HCS and C-HCS. In this section I tested whether the protein:protein interaction between HCS and its protein substrate in the second partial reaction was negatively effected by the mutants using the Yeast Two-hybrid assay (figure 3.8). The positive control here was LexA HCS + B42AD hPC107 which demonstrated that both proteins are correctly folded and functional when expressed as fusion proteins in the Yeast Two-hybrid assay. The negative controls were the self activation control LexA HCS + B42AD (empty) and LexA (empty) + B42AD hPC107. In the dual selection system used, cell growth and blue staining confirmed an interaction and the negative controls showed growth but no staining. White colonies were false positives due to non-specific activation of the LEU2 reporter.

The MCD mutants failed to express lacZ activity, as visualised by a lack of blue staining (figure 3.8) and thereby they disrupted the interaction between HCS and hPC107. The novel mutants identified in this study L166R, L206P, W210R, H306R, F321S HCS also weakened or disrupted the intermolecular interaction. Interestingly, L246M and L270S did not disrupt the intermolecular interaction and therefore must have affected HCS activity via an alternative mechanism. This showed that the inactivation of HCS was due to a different mechanism as the colonies were stained blue. To test whether the disruption of the interaction between HCS and hPC107 was due to poor protein expression or stability Western blot analysis was performed (figure 3.9). The expected size of LexA HCS is 92 kDa and figure 3.9 (A) shows that the fusion proteins were expressed except for the sample that contained LexA H306R HCS (lane 11 in figure 3.9 (A)).

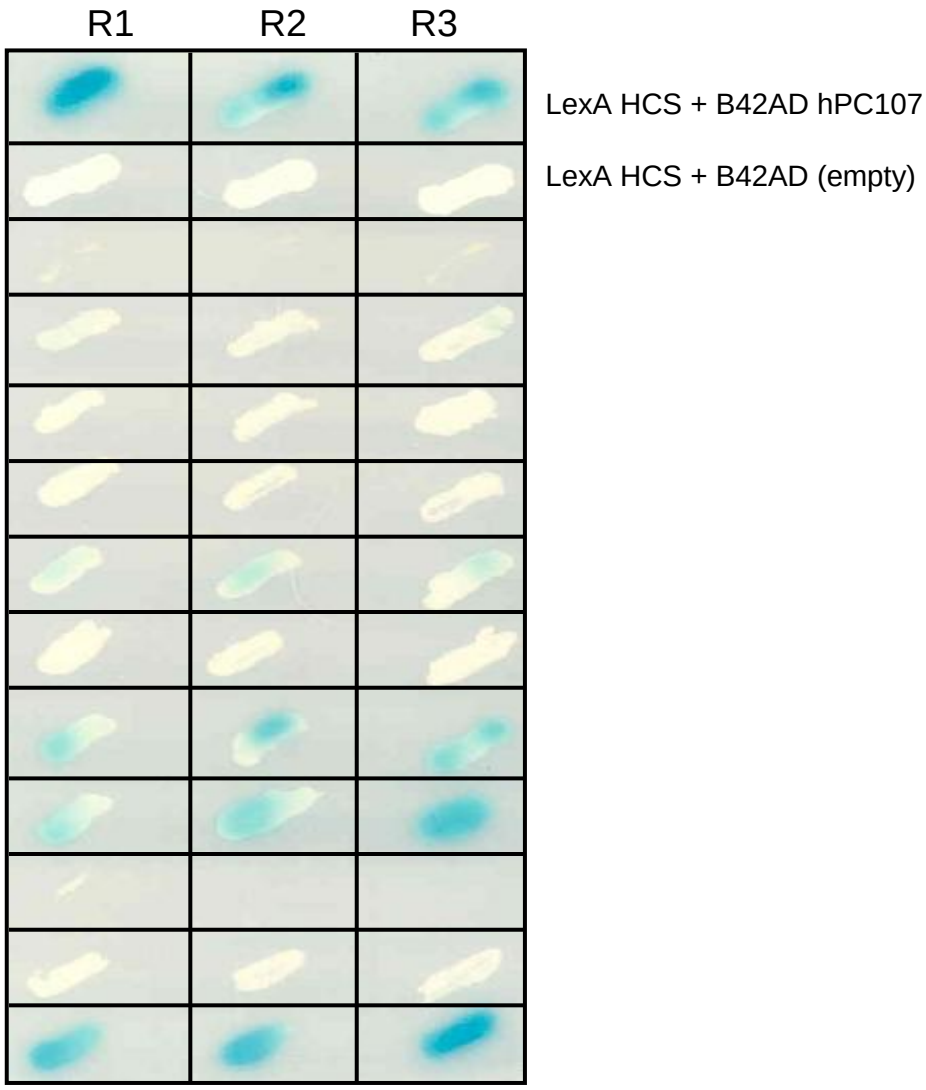


Figure 3.8 Effect of MCD and randomly generated novel HCS mutations on the intermolecular interaction of HCS with hPC107 Yeast Two-hybrid. R1-R3 grid shows three independent clones plated on induction/selection media (galactose, -URA, -HIS, -TRP, -LEU, +X-gal). Growth and blue colour indicates dual activation of LEU2 and LacZ reporter genes.

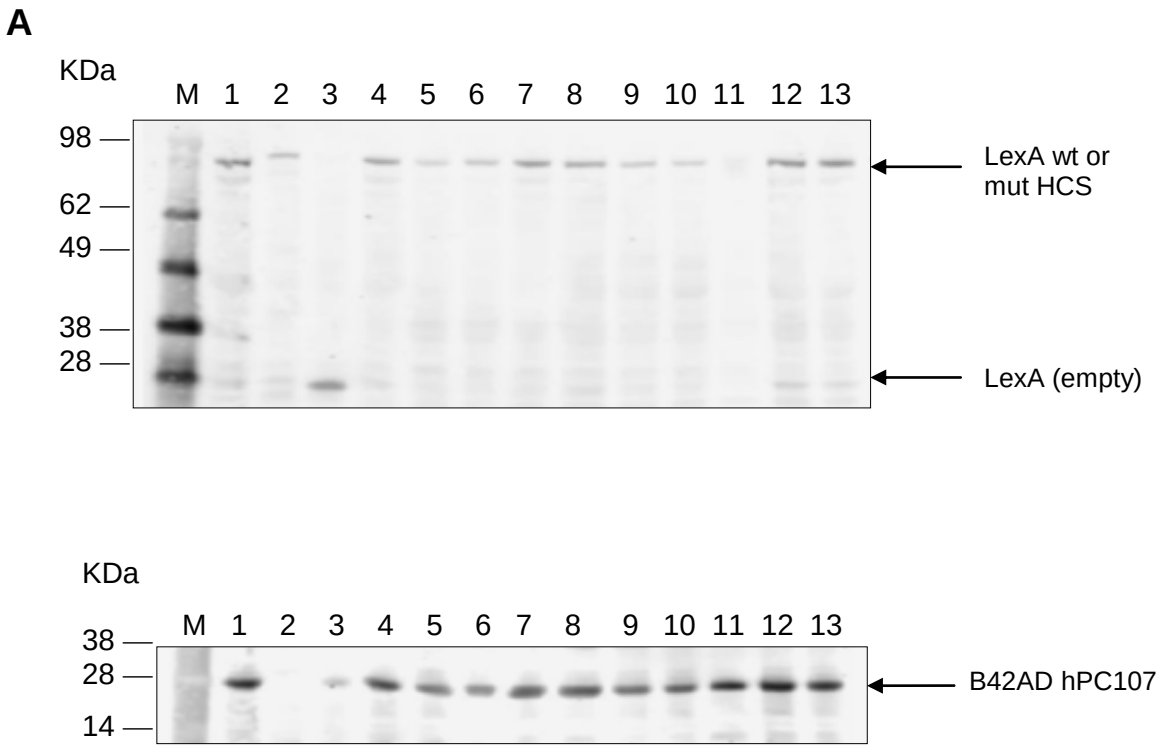


Figure 3.9 Western blot analysis to confirm expression of fusion proteins used in Yeast Two-hybrid assay of intermolecular interaction of wild type and mutant HCS with hPC107.
(A) was probed with rabbit anti-LexA primary antibody and Cy5 anti-rabbit secondary antibody.
(B) was probed with goat anti-HA primary antibody and Cy2 anti-goat secondary antibody. A single membrane was used.
M is Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen)

1 - LexA HCS + B42AD hPC107
2 - LexA HCS + B42AD (empty)
3 - LexA (empty) + B42AD hPC107
4 - LexA HCS [L216R] + B42AD hPC107
5 - LexA HCS [L237P] + B42AD hPC107
6 - LexA HCS [L166R] + B42AD hPC107
7 - LexA HCS [L206P] + B42AD hPC107
8 - LexA HCS [W210R] + B42AD hPC107
9 - LexA HCS [L246M] + B42AD hPC107
10 - LexA HCS [L270S] + B42AD hPC107
11 - LexA HCS [H306R] + B42AD hPC107
12 - LexA HCS [F321S] + B42AD hPC107
13 - LexA HCS [E181G, E327G] + B42AD hPC107

The lack of an interaction between H306R HCS and hPC107 (figure 3.8) was due to a lack of expression of the LexA H306R HCS fusion protein (figure 3.9 (A)). This finding was not surprising as H306R HCS was the only mutation that was unstable or not expressed in section 3.3.5.

3.4 Discussion

From the findings of this study, I conclude that the mechanism by which the MCD mutants affect HCS catalytic activity is by the disruption of the second partial reaction. In order to validate this, biophysical characterisation of the interaction will be presented in Chapter 5.

The current Chapter aimed to characterise the mechanism of biotin non-responsiveness of the V_{MAX} MCD mutants. To address this several studies were conducted. A series of novel HCS mutants were isolated using error prone PCR and an *in vivo* complementation assay. These novel HCS mutants, together with the MCD mutants L216R HCS and L237P HCS, were used to further characterise the function of the N-terminal region of HCS. Using the Yeast Two-hybrid assay, the interaction between HCS and its model substrate hPC107 was characterised by determining whether the novel and MCD HCS mutants disrupt this intermolecular interaction. Most of the novel mutants either weakened or disrupted the interaction. Western blot analysis confirmed that all mutant HCS samples except LexA H306R HCS were expressed. In these yeast cells the lack of an interaction between H306R HCS and hPC107 was most likely due to a lack of expression of the LexA H306R HCS fusion protein. Western blot analysis confirmed that the interaction between HCS and hPC107 was disrupted due to the relevant mutations in HCS and not due to protein degradation, or the lack of expression of the fusion proteins. This data also indicated that the N-terminal region of HCS is involved in protein substrate binding as the interaction between HCS and hPC107 was disrupted by the presence of the mutations in the N-terminal region of HCS. A recent study by Lee *et al.* (2009) showed that residues M1-L160 HCS interacted with hACC75 (a 75 amino acid peptide from human Acetyl CoA Carboxylase) using NMR titration. In another report, N-terminal (residues 1-446) and C-terminal fragments (residues 669-718) of HCS were shown to interact with p67 (a 67 amino acid peptide from propionyl-CoA carboxylase) using a GAL4 Yeast Two-hybrid

assay that uses only growth selection (Hassan *et al.*, 2009). The interaction with 1-446 HCS was qualitatively described as 'weak', given the 21 day incubation period. In this thesis the growth phenotypes were scored after only 3-4 days to minimise the possibility of false positives and appropriate controls were included to address auto-activation of the reporters. These studies combined with the findings of this Chapter indicate the importance of the N-terminal region of HCS in substrate binding.

Although Western blot analysis indicated that the HCS mutants were expressed, the stability of the enzyme has been shown to be affected as determined in the Bailey *et al.* (2008) study, where they characterised the L216R HCS mutant. Bailey *et al.* (2008) showed that the turnover rate for L216R HCS was double that of wild type HCS in mammalian cultured cells. The study also found that while L216R HCS mRNA was detected in patient cell lines, L216R HCS could not be detected, consistent with a high degradation rate.

There is currently no high resolution X-ray crystal structure of HCS. Its elucidation would assist in our understanding of structure and function relationships of HCS. Complexes with ligands and biotin domain substrates would also greatly assist in determining the role of the N-terminal extension. In the absence of a structure, alternative approaches in understanding the N-terminal region of HCS are required. Here I used the Yeast Two-hybrid technique to understand the role of the N-terminal region in HCS function. Another novel finding of this study was the demonstration of a direct interaction between the N-terminal and C-terminal halves of HCS. The interaction between the N-terminal and C-terminal halves of HCS was not affected by the MCD mutants nor the novel HCS mutants. Western blot analysis indicated that the HCS mutants were expressed, except the H306R for which no conclusions could be drawn.

The interaction between the N-terminal and C-terminal halves of HCS was defined by truncation studies to the proposed structured domain encompassing residues 159-314 N-HCS. I propose that this domain may function as a N-terminal cap (NTC) closing over the active site, possibly to help stabilise the labile reaction intermediate from diffusing out of the active site. It could not be concluded, however, whether 1-151 N-HCS was able to interact with the C-terminal region of HCS as Western blot analysis showed no evidence of expression. Lee *et al*. (2009) were able to express (in *E. coli*) and purify 1-160 N-HCS. Lee *et al.* (2009) showed that the predicted structure of this region was mostly random coil. This may explain the degradation or lack of expression that was observed in the yeast system as unstructured regions would be prone to proteolysis. Phyre was used to predict the structure of NTC and it was predicted to be structured and this will be discussed in more detail in Chapter 4. The data generated in this study gives an indication that the N-terminal region of HCS (particularly NTC) may have a role in catalysis through its interaction with the C-terminal or catalytic region of HCS.

The results in this Chapter give an indication that in MCD V_{MAX} patients the mutations affect the interaction between HCS and its substrate. It is not the interaction between the N-terminal and C-terminal halves of HCS that are disrupted by the mutations. This Chapter showed that the function of NTC could be substrate binding and that it interacts with the C-terminal region of HCS. I have shown in this Chapter that 159-314 N-HCS interacts with C-HCS and in Chapter 4 the determination of the structure of 159-314 N-HCS will be discussed. Alternative approaches to validate the Yeast Two-hybrid results are required. In Chapter 5 several approaches were pursued including GST pull-downs and surface plasmon resonance (SPR).

CHAPTER 4

NMR studies of the N-terminal cap (NTC) of HCS

4.1 Introduction

As discussed in Chapter 1, X-ray crystal structures have been determined for class I and II BPLs from archea (Bagautdinov *et al.*, 2005) to bacteria (Wilson *et al.*, 1992; Pendini *et al.*, 2008). A high-resolution structure of a large eukaryotic BPL, including HCS, has yet to be elucidated. Due to difficulties in expressing large quantities of full length HCS (Chapter 5) and difficulties in expressing the C-terminal region of HCS (Chapter 5), the N-terminal cap (NTC) was selected as a candidate for obtaining structural information on HCS. In Chapter 3, it was shown that the NTC interacts with the C-terminal half of HCS and that it is also involved in substrate recognition. Class I and II BPLs do not contain the large N-terminal extension contained in HCS and therefore obtaining structural information on a class III is important for a greater understanding of this class III BPL.

4.1.1 Structural Determination of proteins by NMR Spectroscopy

The two main techniques for obtaining atomic resolution structures of proteins are X-ray crystallography and solution nuclear magnetic resonance (NMR) spectroscopy. X-ray crystallography has been valuable for expanding our knowledge of protein structures. However, the major limitation with this technique is that it requires the formation of protein crystals. NMR is an alternative approach that can be used to determine the three-dimensional structures of biological macromolecules in solution thereby avoiding the need for crystallography. NMR lends itself favourably to protein up to 30-40 kDa (Dotsch and Wagner, 1998).

In this study, NMR spectroscopy was employed. NMR spectroscopy uses the magnetic spin properties of atomic nuclei, such as ^{1}H, ^{15}N, and ^{13}C, to identify physically close atoms (Howard, 1998). When oriented by a magnetic field, the nuclei absorb radiation at a particular frequency depending on their local molecular environment. The frequency at

which nuclei resonate is measured as chemical shift in parts per million (ppm) relative to a reference signal. In a one-dimensional (1D) ^{1}H spectrum the intensity of the signal is measured and plotted against chemical shift (Roberts, 1993).

In NMR, protons contained within similar molecular structures, such as functional groups, resonate at similar frequencies. When the protein is unstructured, protons in the following functional groups resonate at characteristic frequencies: methyl group (CH_3) at ~1 ppm; methylene group (CH_2) at ~2-3 ppm; α-protons at 4-5 ppm; aromatic groups at 6-7.5 ppm; and, finally, amide groups at 7-11 ppm (Wuthrich, 1986). Evidence of folding is determined when the frequency of proton resonance is shifted up field (lower ppm value) or downfield when the polypeptide chain is folded and contacts are formed between different residues.

Spectral overlap is observed when several protons exist in similar environments. In order for full protein structures to be determined by NMR 1D spectra need to be extended into two or more dimensions. This is because 1D NMR spectra of biological macromolecules usually contain hundreds of resonance peaks which overlap heavily and cannot be interpreted. In 2D experiments, the interaction between the same nuclei (homonuclear i.e. ^{1}H-^{1}H) or different nuclei (heteronuclear eg. ^{1}H-^{15}N) are observed. These experiments include correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY) and nuclear Overhauser enhancement spectroscopy (NOESY) experiments. COSY and TOCSY experiments reveal information about adjacent nuclei that have through-bond linkages. NOESY experiments detect interactions between nuclei that are close together in space (less than 5 Å apart) but not necessarily close in the linear sequence of the peptide (Wuthrich, 1986). In this study TOCSY and NOESY experiments were conducted.

To resolve a protein structure, each resonance needs to be mapped to its corresponding nucleus using sequence-specific resonance assignments. After building a list of distance restraints between nuclei, the protein structure is resolved by creating a model in which each distance restraint is satisfied. To achieve this the steps required are preparation of the protein solution, recording NMR spectra, assignment of NMR signals to individual atoms in the molecule, identification of conformational restraints, such as distance restraints, torsion angle restraints and orientational restraints, and finally calculation of the three-dimensional structure (Spronk *et al.*, 2004).

The aim of this Chapter was to determine the structure of the N-terminal cap (NTC) using NMR.

4.2 Specific Methods

4.2.1 Cloning of pGEX NTC, the NTC variants pGEX NTC (G159-E293) and pGEX NTC (G159-D349)

In section 3.2.1 the pB42AD NTC construct was made. The DNA of this construct was used for the subcloning of DNA encoding NTC into pGEX-4T. pB42AD NTC and pGEX-4T-2 were digested with *Eco*RI and *Xho*I (as described in Chapter 2, section 2.2.5.7), an agarose gel was run and DNA encoding NTC and digested pGEX-4T were excised from the gel. DNA encoding NTC and pGEX-4T were purified using the GeI purification kit (QIAGEN) and each insert was ligated into pGEX-4T using DNA ligase (Chapter 2, section 2.2.5.8). The ligation was transformed into competent *E. coli* DH5α cells and the clones were analysed by restriction digest with *EcoR*I and *Xho*I. Clones with the correct sized fragments were confirmed by DNA sequencing using pGEX forward and reverse primers. To ensure that the fusion protein was in frame, phosporylation and annealing of B225 and B226 oligonucleotides was performed as decribed in Chapter 2, section 2.2.7.4. The oligonucleotides also contained an *Eco*RV site, which was a unique restriction site in order to facilitate screening. The annealed oligonucleotides were ligated into each digested construct and the ligations were transformed into competent *E. coli* DH5α cells and the clones were analysed by restriction digest with *Eco*RV. Clones with the correct sized fragments were confirmed by DNA sequencing using the pGEX forward and reverse primers. pGEX NTC (G159-E293) and pGEX NTC (G159-D349) were cloned using pET HCS-His6 as the DNA template for PCR. The oligonucleotides that were employed were B277 and B278 for pGEX NTC (G159-E293), B277 and B279 for pGEX NTC (G159-D349). B277 contained a *Bam*HI site at the 5´ end and B278 and B279 contained a *Xho*I site at the 3´ end. Using *Bam*HI ensured that the proteins were in frame. Purified PCR products were digested with *Bam*HI and *XhoI* and ligated into *Bam*HI and *XhoI*-treated pGEX-4T-2 plasmid. All the ligations were transformed into competent *E. coli* DH5α cells and the

clones were analysed by restriction digest with *Bam*HI and *Xho*I. Clones with the correct sized fragments were confirmed by DNA sequencing using the pGEX forward and reverse primers.

4.2.2 Transformation, expression and purification of GST-NTC (G159-R314), GST-NTC_2 (G159-E293) and GST-NTC_3 (G159-D349)

The residues G159-E293 and G159-D349 were chosen based on the sequence alignment of the N-terminal region of HCS against other vertebrate species (Chapter 3, figure 3.6). These residues were chosen in order to improve upon the protein yield of NTC. pGEX NTC (G159-R314), pGEX NTC_2 (G159-E293) and pGEX NTC_3 (G159-D349) were transformed into *E. coli* BL21 cells. Overnight samples of *E. coli* BL21 containing each pGEX NTC in LB culture (containing 0.1 mg/mL ampicillin) were sub-cultured to 1 L by diluting 1:50 and were incubated at 30°C until it reached an OD_{600} of 0.6. Induction of the GST-NTC proteins was performed by adding 0.1 mM IPTG and the culture was incubated a further 3 hours. Cells were harvested and resuspended in 20 ml of lysis buffer (Tris buffered saline, pH8.5) per 1 L culture and lysed with a cell disruptor. The lysate was pelleted by centrifugation at 21 000 x g for 20 min. The supernatant was filtered using a 0.8 µm followed by a 0.45 µm filter.

4.2.3 Cloning, transformation and expression of NTC using alternative expression vectors

Shee Chee Ong (University of Adelaide) cloned pET32a NTC and pYEX NTC. These constructs were transformed into *E. coli* (DE3) and *S. cereviseae* w303, respectively.

4.2.4 Purification of NTC using on column cleavage

Expression of GST-NTC (G159-R314) and preparation was performed as in section 4.2.2. Purification of NTC was performed using 10 mL glutathione agarose beads packed in a column (sample 1). This was performed by equilibrating the column with TBS, pH 8.5. The supernatant was applied two times through the column. The glutathione agarose beads were washed with 10 column volumes of TBS, pH 8.5 or until no blue colour was detected by the Bradford assay. On column cleavage was performed by adding 10ml of TBS, pH 8.5 containing 30 U of thrombin and 2.5 mM $CaCl_2$. Thrombin cleavage was performed at 37°C for 2 hours followed by the elution of cleaved NTC using TBS, pH 8.5. Glutathione-bound GST was eluted using TBS, pH 8.5 containing 10 mM glutathione. Where the purification of NTC was performed using the Profinia^{TM} Protein Purification System (BIO-RAD) (sample 2) this was performed as in Chapter 2 section 2.2.2. The sample was buffer exchanged in 10 mM phosphate buffer, pH 6.8 using a PD10 column and stored at 4°C. Glutathione-bound GST was eluted using TBS, pH 8.5 containing 10 mM glutathione. SDS-PAGE was employed to analyse the purifications. Due to the low yield of NTC purified, 4 independent purifications from 2 L cultures each were performed and pooled together for each NTC purification method. The samples were stored at 4°C until the NMR experiments were performed. Before performing the NMR experiments 10 % v/v deuterium oxide (D_2O) was added to the sample. In NMR D_2O provides the lock signal, which prevents random drift of the magnetic field relative to a reference frame.

4.2.5 2D NMR Spectroscopy Experiments

Preliminary NMR experiments were conducted using a Varian Inova 600 spectrometer using the BioPack package in the Department of Chemistry, Adelaide University with the assistance of Associate Professor Grant Booker. All data sets were recorded at 25°C using a 5 mm inverse triple resonance $^{1}H/^{13}C/^{15}N$ PFG probe in the first experiment and the

$^{1}H/^{13}C/^{15}N$ PFG cold probe in the second experiment. The carrier frequency was centred on the H_2O signal. Two dimensional homonuclear experiments were recorded. These included 2D ^{1}H TOCSY and ^{1}H NOESY experiments. The NMR data sets were recorded and processed using the VNMR software. ^{1}H NOESY experiments reveal information about through-space interactions between different protons. In these 2D contour plots, each proton is represented as a peak on the diagonal. A cross peak (off the diagonal) is observed when two protons are within approximately 5Å of each other and this can be extrapolated to the diagonal to identify the interacting protons (Clowes *et al.*, 1995).

4.2.6 Investigation of the secondary structure composition, structural integrity and thermal stability of NTC using Circular dichroism

NTC was purified again as in section 4.2.4 (sample 3) and purified NTC was buffer exchanged by dialysis into 10 mM sodium phosphate buffer, pH 8.0, using regenerated cellulose tubular membrane (Fisher Biotech, Australia). Mass spectrometry (MS) analysis of tryptic digestion products of purified NTC and mass determination of NTC using MS was performed.

The pure NTC sample was diluted and the concentration was determined to be 0.25 mg/mL using the Bradford assay and separated into two for the different experiments. CD was carried out on NTC at 20°C as described in Chapter 2, section 2.2.6. In order to determine whether NTC unfolds at 90°C and whether this unfolding is reversible, CD was performed on the same sample at 90°C and then again at 20°C. All spectra were recorded with an average of 5 scans with subtraction of the buffer blank 10 mM sodium phosphate buffer, pH 8.0. The generated data was used to determine the wavelength for monitoring thermal stability. A melting experiment was performed on the second NTC sample using CD by monitoring the ellipticity at 220nm at every 1°C ranging from 20°C to 90°C. From this data the melting temperature (T_m) was determined.

Dichroweb (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml: Whitmore and Wallace, 2008) was used to analyse the NTC spectrum using the programs CDSSTR, CONTINLL and SELCON3 at a reference set optimised for 185 to 240 nm.

4.2.7 Investigation of the oligomeric state of NTC using gel filtration chromotography

NTC that was purified in 4.2.6 was used for the determination of the oligomeric state of NTC. To obtain the molecular mass of NTC in solution, size-exclusion chromatography with on-line light scattering, absorbance, and refractive index detectors were used. Size-exclusion chromatography was carried out with a Pharmacia ÄKTApurifier UPC 10 (General Electric Company). A Superdex^{TM} 200 10/300GL column was connected in-line to a UV detector ÄKTA pH/C-900 (GE Company), a miniDAWN TREOS (Wyatt Technology), multi-angle laser light-scattering detector and an Optilab rEX (Wyatt Technology) refractive index detector. Both the light scattering detector and the refractive index detector have a heating/cooling system that controlled the temperature at 25°C. A 100 µl sample (10 mg/ml BSA followed by 1 mg/ml NTC in a separate experiment) was injected and the column was eluted with 10 mM Tris, pH 8.0 and 150 mM NaCl buffer at a rate of 0.25 ml/min. The molecular mass was calculated using the ASTRA V 5.3.2.21 software (Wyatt Technology) using the Debye fitting method (Folta-Stogniew and Williams, 1999).

4.2.8 Molecular structural models of 159-314 N-HCS or NTC and 1-151 N-HCS

The ExPasy Proteomics Tools server (http://au.expasy.org/tools/) provided access to tertiary protein structure prediction program (Phyre, automated 3D model builder which uses profile-profile matching and secondary structure). Phyre version 0.2 was used to predict the structure of NTC or 159-314 N-HCS and 1-151 N-HCS. The amino acid sequence for each protein was entered into the program and the resulting structure was based on homology with proteins in the data bank.

4.3 Results and Discussion

4.3.1 Selection of NTC construct

Several constructs of NTC were made and transformed into the appropriate bacterial or yeast host. Multiple constructs were investigated as the exact boundaries of NTC are yet to be confirmed by structural biology. Small scale expression studies were performed to investigate whether the proteins could be expressed as a soluble protein. The pGEX NTC_2 (G159-E293) and pGEX NTC_3 (G159-D349) constructs were chosen based on the sequence alignment of N-terminal HCS against other eukaryotic species (Chapter 3, figure 3.5). The domain boundaries were either before or after the R314 residue. Although these GST fusion proteins could be expressed, they were not soluble even after varying growth temperatures, incubation times and IPTG concentrations. The domain boundaries of pGEX NTC (G159-R314) were determined by limited proteolysis of the N-terminal region of HCS, performed previously in our lab by Rachel Swift (2004). Importantly, this was the only construct that was produced as a soluble protein (table 4.1). Hence GST-NTC (G159-R314) was selected for subsequent investigation.

Table 4.1 Selection of a suitable construct for the expression and purification of NTC

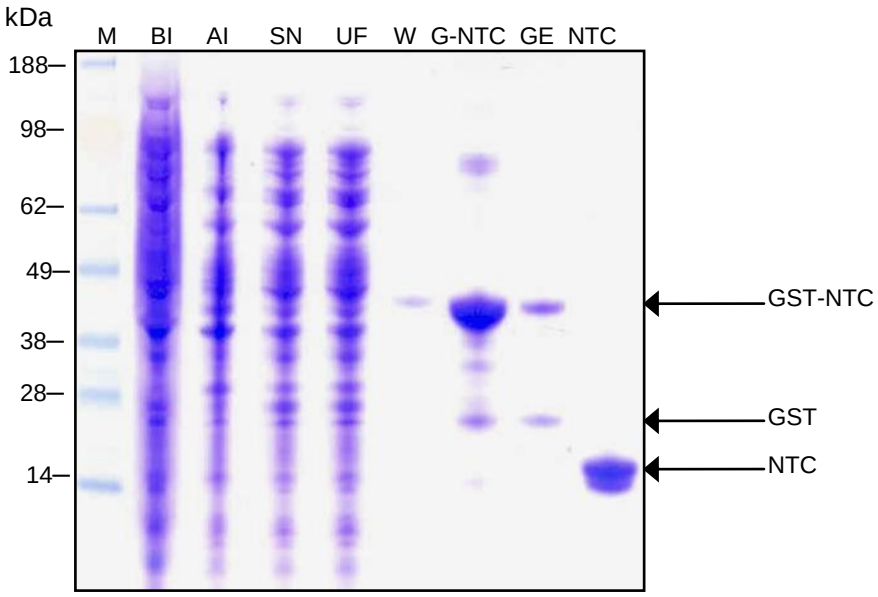
Construct	Transformed	Expressed	Soluble
pGEX NTC (G159-R314)	Yes	Yes	Yes
pET32a NTC (G159-R314)	Yes	No	N/A
pYEX NTC (G159-R314)	Yes	No	N/A
pGEX NTC_2 (G159-E293)	Yes	Yes	No
pGEX NTC_3 (G159-D349)	Yes	Yes	No

4.3.2 Expression and purification of NTC (sample 1)

E. coli BL21 containing pGEX-NTC was used to express GST-NTC. Purification of the NTC was performed using 10 mL glutathione agarose beads packed in a column. The purification method was outlined in section 4.2.4. Samples from the various stages of the NTC purification were analysed using SDS-PAGE (figure 4.1A). A protein around the expected size of NTC (18.3 kDa) was observed following thrombin cleavage. It was estimated that NTC was >90% pure. Protein yield was calculated using the Bradford assay to be 2.4-2.8 mg total NTC from 2 L culture. Four purifications of NTC (8 L total of cultured cells) were pooled and concentrated to 0.52 mM using a 3000 MWCO column (Amicon). No precipitation was observed.

The protein purification was again repeated, this time using the Profinia™ Protein Purification System (BIO-RAD) using the Bio-Scale™ Mini Profinity™ GST 5 ml cartridge (figure 4.1B). In this second purification, 2.1-2.5 mg total NTC from 2 L culture and the four purifications of NTC were pooled and concentrated to 0.5 mM for NMR analysis. It was estimated that NTC was >90% pure. No precipitation was observed. Both samples 1 and 2 were used for NMR analysis (section 4.3.5).

A



B

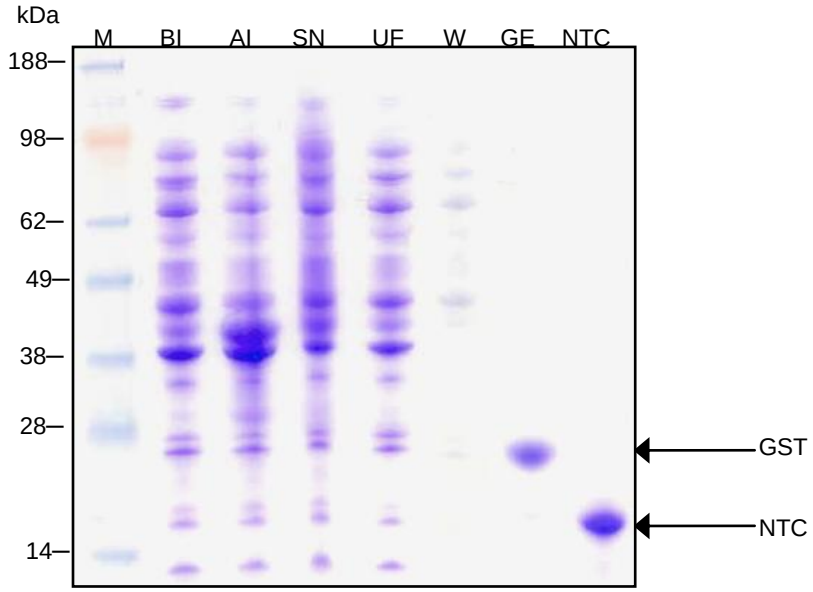


Figure 4.1 Purification of NTC HCS **(A)** using a glutathione affinity chromatography. M - Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen), BI - before induction, AI - after induction, SN - supernatant, UF - unbound fraction, W - wash, G-NTC - GST-fused NTC, GE – Glutathione elution, NTC - concentrated fraction from 4 purifications of 2 L each.
(B) Purification using the Profinia™ Bio-Scale™ Mini Profinity™ GST cartridge. M - Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen), BI - before induction, AI - after induction, SN - supernatant, UF - unbound fraction, W - wash, GE - Glutathione elution, NTC - concentrated fraction from 4 purifications of 2 L each.

4.3.4 Expression and purification of NTC (sample 3), indentification of NTC using Mass Spectrometry (MS) and mass determination of NTC using MALDI MS

Finally a third preparation of NTC was performed. This time a 50 mL glutathione-agarose column was employed. This time 2.8 mg total protein was obtained at >90% purity (figure 4.2) and no precipitation was observed. The protein from this preparation was used for confirmation by mass spectrometry, for gel filtration chromatography and CD analysis.

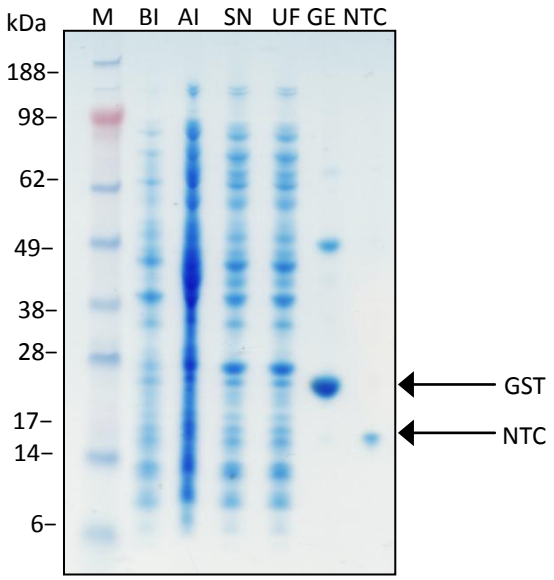


Figure 4.2 Purification of NTC HCS. M - Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen), BI - before induction, AI - after induction, SN - supernatant, UF - unbound fraction, GE - Glutathione elution and NTC

NTC was subjected to in-gel tryptic digest and analysis by liquid chromatography-electrospray ionisation tandem mass spectrometry (LC-ESI-MS/MS) to confirm the identity of this protein. Peptides with multiple-charged ions were automatically detected and subjected to fragmentation. Fragmentation data were used to search the MASCOT database for protein identities. All 30 peptides that were identified matched the NTC sequence with 64% coverage (figure 4.3). This confirmed, with greater than 95% confidence, that the purified species was indeed NTC.

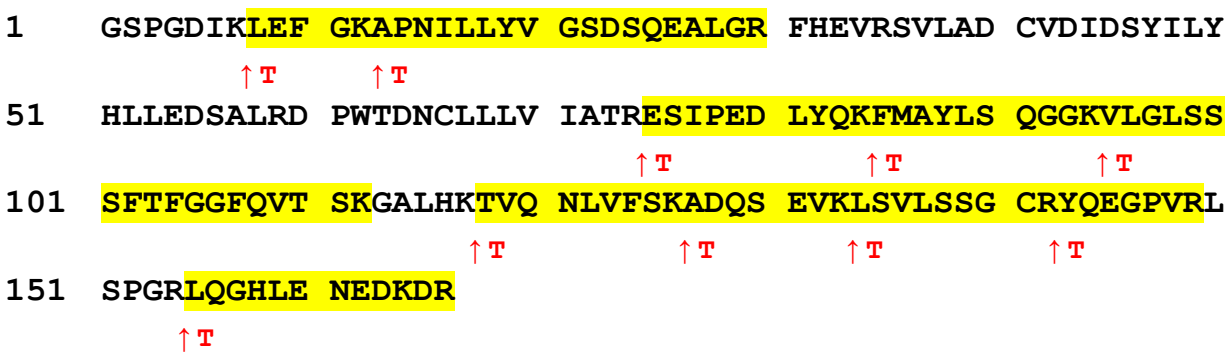


Figure 4.3 LC-ESI MS/MS analysis of NTC
NTC sequence showing trypic peptides identified in the spectra and trypsin cleavage sites ↑T

NTC is predicted to be 17266.5 Da. Due to the cloning strategy described in section 4.2, the actual size of the NTC in this study is 18310.7 Da. Figure 4.4 shows how the introduction of the EcoRV cloning site at the 5` end of the NTC coding region extends the N-terminal region of NTC with a GSPGDIKLEF motif.

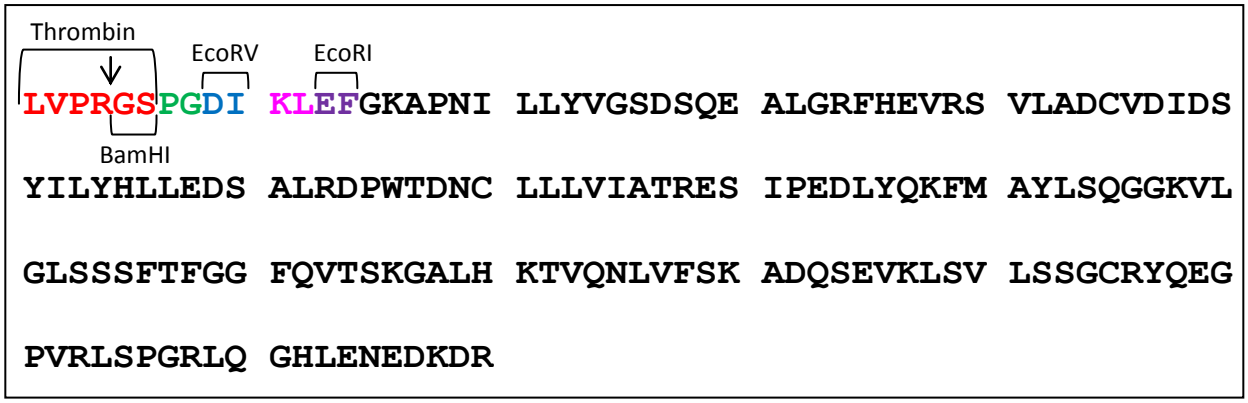


Figure 4.4 Amino acid sequence of NTC
Thrombin cleavage site, the introduced EcoRV site, and the EcoRI and BamHI sites are shown. KL, depicted in pink are the resulting amino acids from inserted nucleotides used to ensure NTC was in frame. NTC is depicted in bold and in black

The mass of the NTC (sample 3) was determined using MALDI MS. Mass spectra were acquired using a MALDI-TOF/TOF mass spectrometer and the mass of NTC determined to be 18341.7 Da (figure 4.5), which is 31 Da larger than the predicted size of NTC. In linear mode MALDI MS has an accuracy of +/- 2 kDa and therefore the size difference was likely 32 Da. The 32 Da size difference was most likely due to post-transional modification of the two cysteins in NTC as the thiol group of cysteine is nucleophilic and eaily oxidised (oxygen atom is 16 Da in size).

Together the LC-ESI MS/MS and MALDI-TOF analysis of NTC confirmed the correct protein was purified.

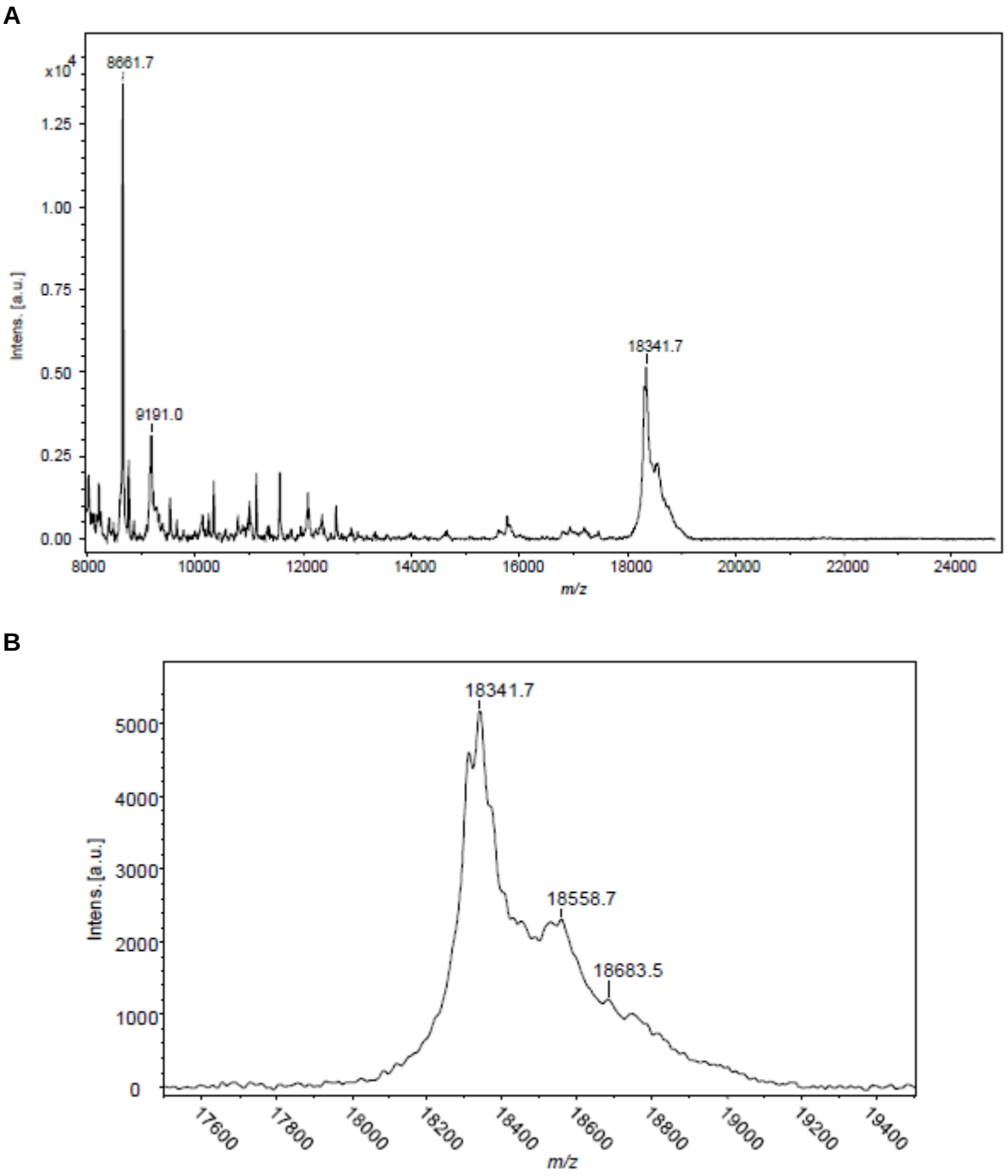


Figure 4.5 MALDI mass spectra of NTC
(A) Positive ion mode, 800-25000 m/z, **(B)** 17500-19500 m/z

The base peak in the positive ion mode spectrum is 8661.7 m/z, as shown in figure 4.5A, which represents the full mass spectrum. A zoomed spectrum covering the m/z range 17500 – 19500 Da is shown in figure 4.5B.

4.3.5 2D NMR experiments of NTC using a triple resonance $^{1}H/^{13}C/^{15}N$ PFG probe (sample 1) and triple resonance $^{1}H/^{13}C/^{15}N$ PFG cold probe (sample 2)

Figures 4.6 and 4.7 show the over-laid TOCSY (blue) and NOESY (red) 2D experiments for samples 1 (0.52 mM) and 2 (0.5 mM), respectively. The triple resonance $^{1}H/^{13}C/^{15}N$ PFG cold probe is more sensitive than the triple resonance $^{1}H/^{13}C/^{15}N$ PFG probe and this is evident in the resulting spectrum (figure 4.7) as there are more peaks. Both spectra are dominated by a small number of sharp, high intensity signals. However, there were very few cross peaks in the NOESY experiment at the amide region (between 0-4 ω_1 - ^{1}H ppm and 4-8 ω_2 - ^{1}H ppm). Figure 4.8 is an example of an ideal 2D ^{1}H-^{1}H NOESY spectrum of α-actinin-4 CH domains performed previously in our lab (Chen, 2008). In this spectrum there were many NOESY cross peaks in the amide region (between 0-5 ω_1 - ^{1}H ppm and 5-10 ω_2 - ^{1}H ppm). The lack of many cross peaks may be due to line broadening and the observed signal may be due to highly mobile residues. This observation indicates that NTC may self-associate. Precipitation of NTC was observed after the experiments were conducted. Line broadening was evident in the two independent experiments, demonstrating the reproducibility of this data.

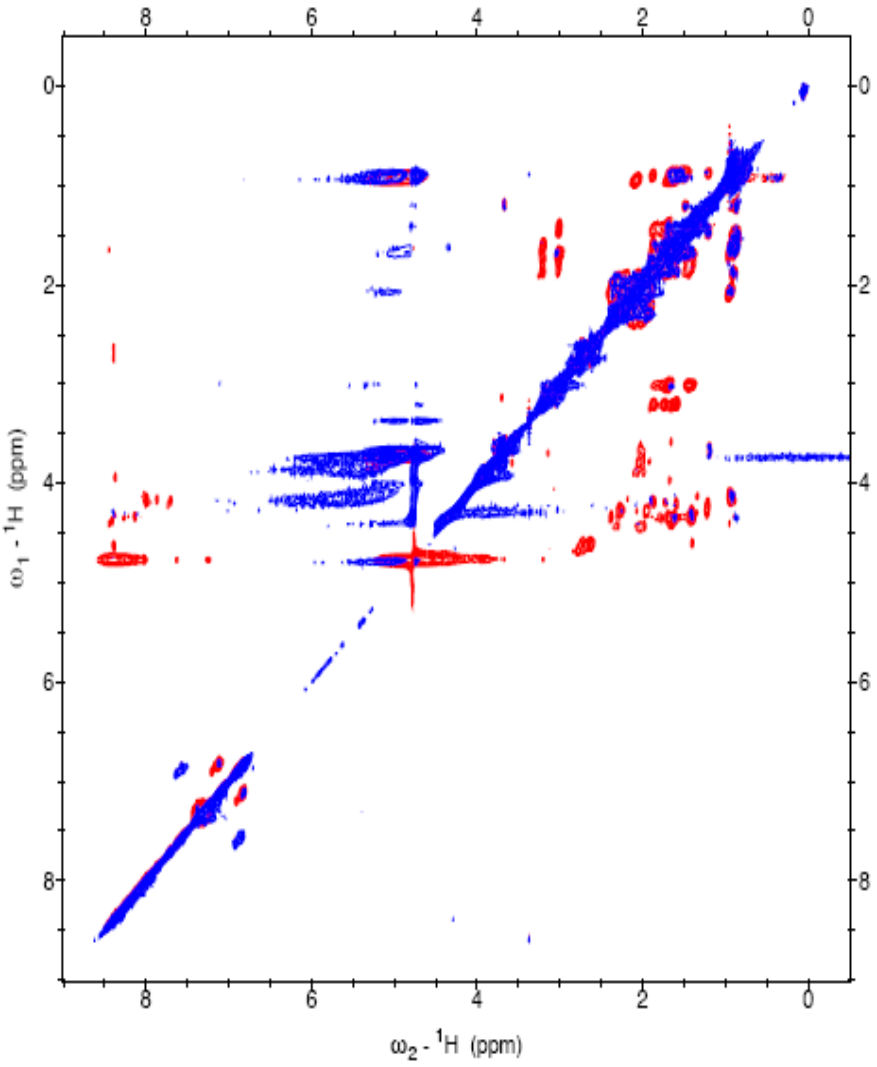


Figure 4.6 Spectra of over-laid TOCSY (blue) and NOESY (red) 2D experiments for NTC sample 1. The spectra were recorded using a 0.52 mM NMR sample at 25˚C. The NMR data were processed by VNMR software and CCPNMR software. The spectra were plotted from CCPNMR showing the region from 0 to 8 ppm in both dimensions.

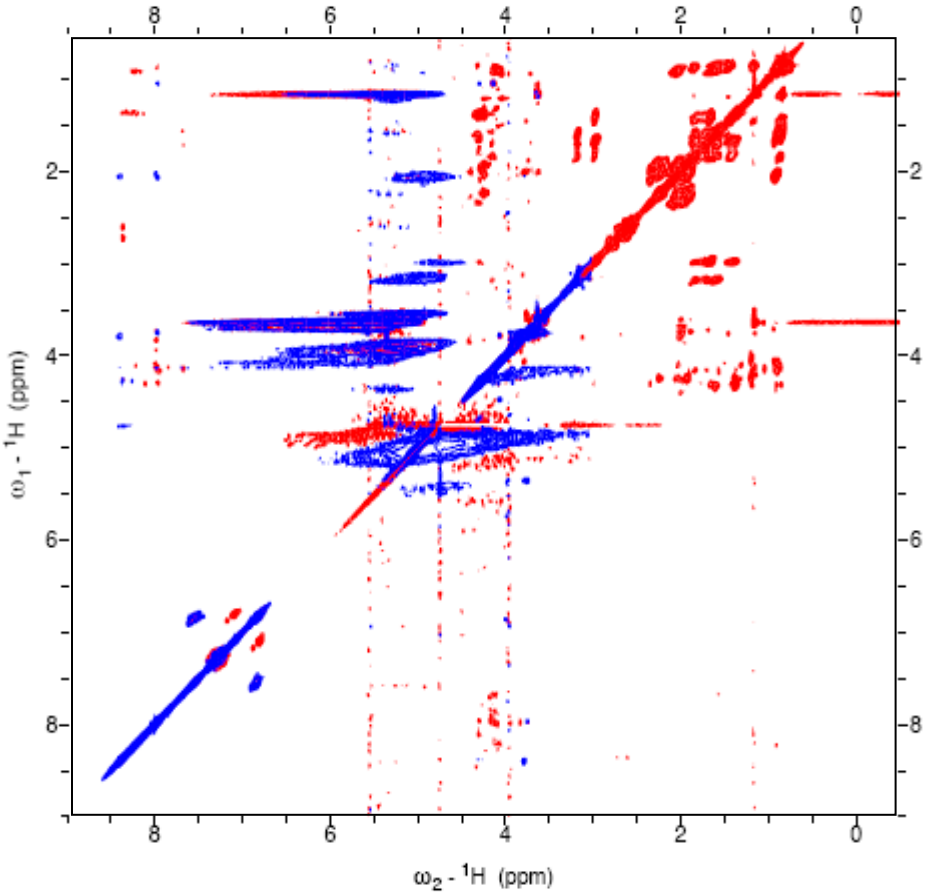


Figure 4.7 Spectrum of over-laid TOCSY (blue) and NOESY (red) 2D experiments for NTC sample 2 (with cold probe). The spectra were recorded using a 0.5 mM NMR sample at 25˚C. The NMR data were processed by VNMR software and CCPNMR software. The spectra were plotted from CCPNMR showing the region from 0 to 8 ppm in both dimensions.

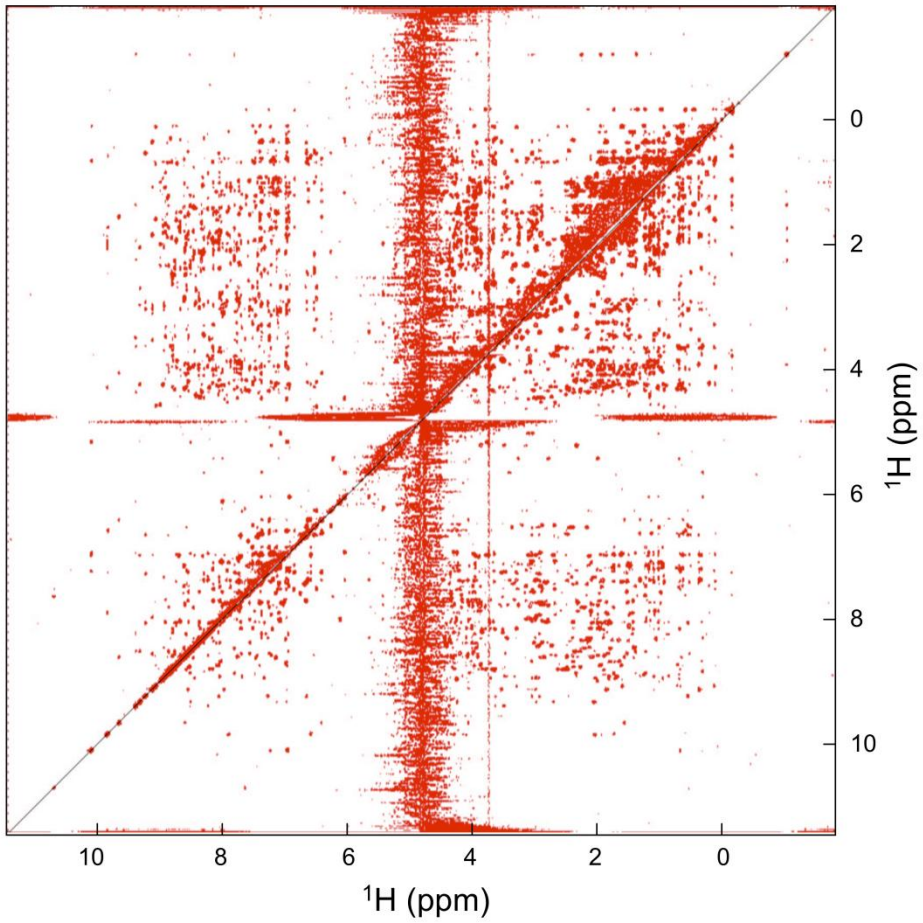


Figure 4.8 2D ^{1}H-^{1}H NOESY spectrum of α-actinin-4 CH domain (Chen, 2008)
The 2D NOESY spectrum was recorded using a 3 mM NMR sample at 25˚C. The NMR data were processed by NMRPipe and analysed by CCPNMR software. The spectrum was plotted from CCPNMR showing the region from 0 to 10 ppm in both dimensions.

4.3.6 Circular Dichroism (CD)

4.3.6.1 CD analysis of NTC

CD is a method that determines the secondary structure of proteins. Typically, α-helical proteins have negative readings at 222 nm and 208 nm and a positive reading at 193 nm (Holzwarth and Doty, 1965). Proteins with antiparallel β-sheets have negative readings at 218 nm and a positive reading at 195 nm (Greenfield and Fasman, 1969) whereas disordered proteins have very low ellipticity above 210 nm and negative readings near 195 nm (Venyaminov *et al.*, 1993). The CD spectrum of 0.25 mg/mL NTC (figure 4.9) exhibited a spectrum typical of a structured α-helical protein. This was in agreement with further analysis of the CD data (see section 4.3.6.2). Performing CD at 90°C in order to denature the protein showed that NTC was folded as the higher temperature induced it to be less structured. The CD spectrum was performed at 20°C after melting and showed that the unfolding of NTC was not completely reversible because the spectrum did not return to the original NTC spectrum obtained before melting (figure 4.9).

The wavelength used to monitor thermal stability was 220 nm as the greatest change in ellipticity was observed at this wavelength. A melting experiment was then performed using CD by monitoring the ellipticity at 220nm at every 1°C ranging from 20°C to 90°C (figure 4.10). From this data the melting temperature (T_m) was determined to be 55.7°C and the R square value was 0.98 using the Boltzmann sigmoidal curve (GraphPad, Prism). The R^2 value is computed from the sum of the squares of the distances of the points from the best-fit curve determined by non-linear regression and the R^2 value of 0.98 indicated that the data fit the curve well (R square value of 1 indicates best fit). No precipitation was observed possibly due to the low concentration used and the fact that the experiment runs for 1 hour. This may explain why the time-dependent precipitation that was observed in the NMR experiments was not evident in the CD experiments.

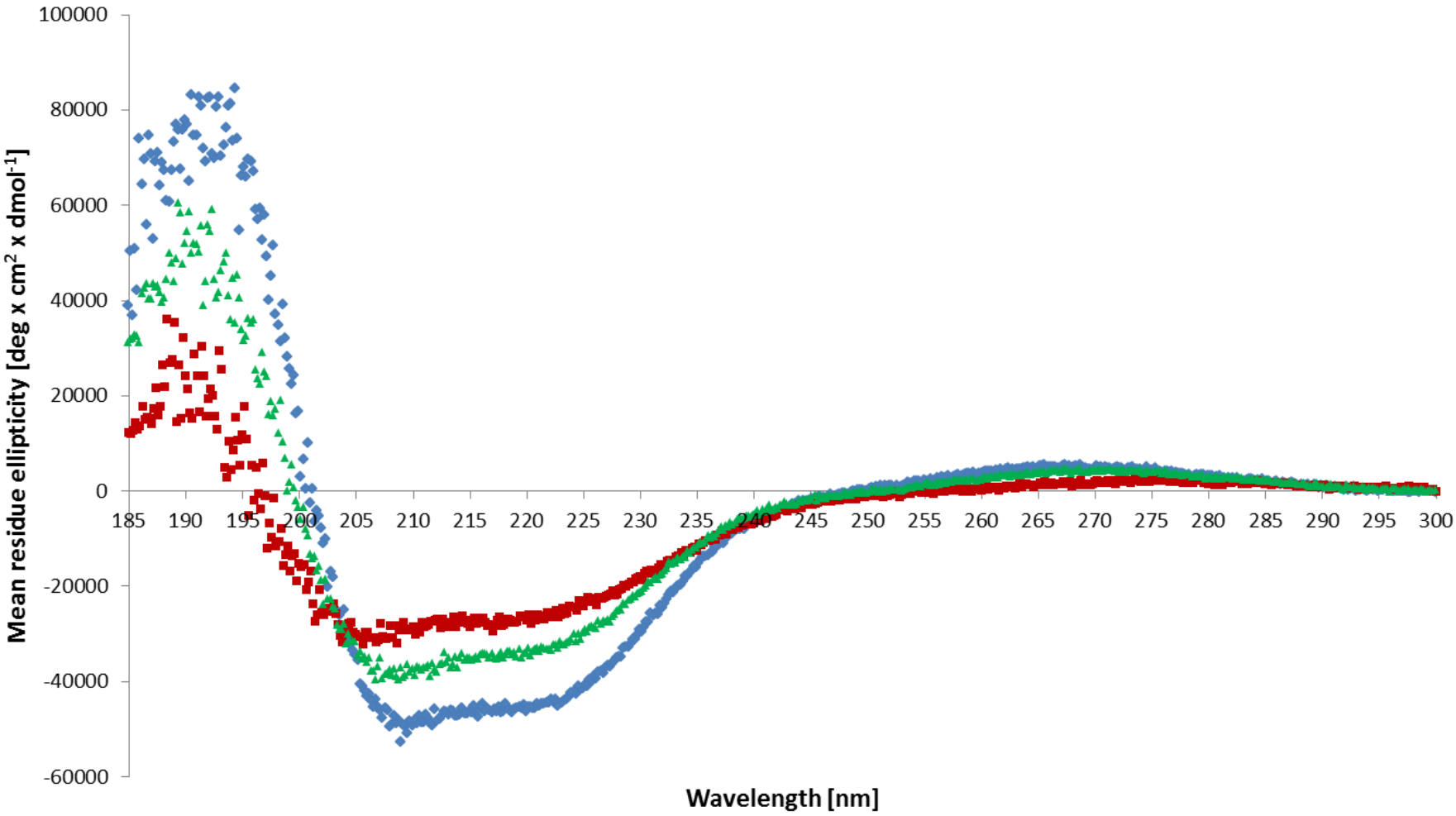


Figure 4.9 Circular Dichroism spectrum of NTC sample.

The NTC sample was exchanged into 10 mM sodium phosphate buffer, pH 8.0 by dialysis. The sample was diluted to 0.25 mg/mL. Individual spectra (including 10 mM sodium phosphate buffer, pH 8.0, as a blank) were recorded at 20°C ◆, then 90°C ■ and back to 20°C ▲ across wavelengths 185 nm to 300 nm with a nitrogen flow of 7.5 L/min, scan speed of 20 nm/min and accumulation mode of 5 (average of 5 scans).

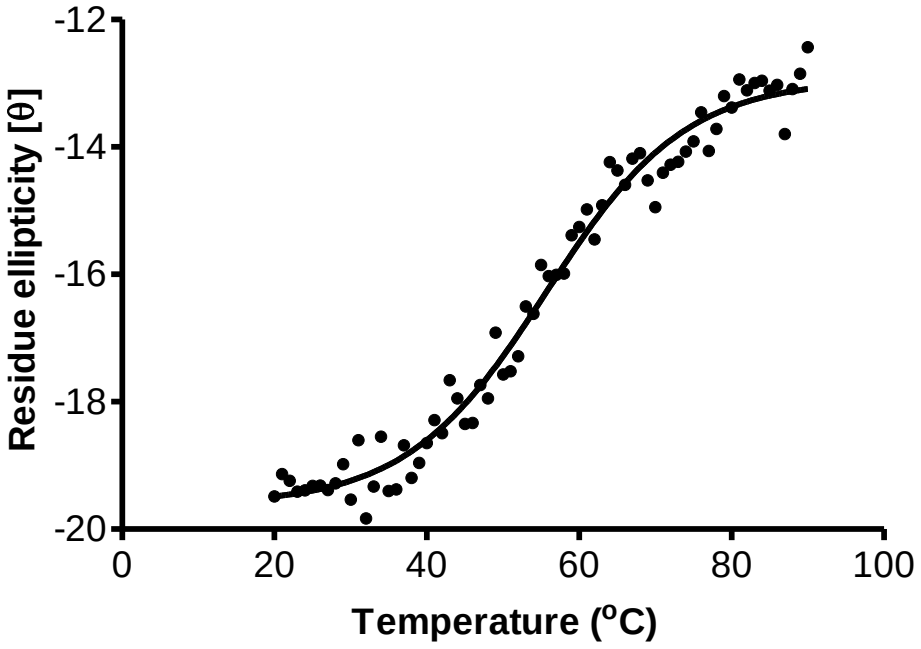


Fig 4.10 The analysis of the melting of purified protein from CD data using the Boltzmann sigmoidal curve

4.3.6.2 Analysis of NTC CD spectrum using CONTINLL, CDSSTR and SELCON3

The programs CDSSTR, CONTINLL and SELCON3, using DICHROWEB (Whitmore and Wallace, 2008), were used to estimate the secondary structure composition of NTC (table 4.2). Here, the NTC CD spectrum used was data at 20°C, before melting (figure 4.9).

Table 4.2 Analysis of NTC spectrum using various analysis programs (on Dichroweb) at a reference set optimised for 185 to 240 nm

Method	Helix 1	Helix 2	Strand 1	Strand 2	Turn	Unordered
CDSSTR	0.520	0.270	0.030	0.060	0.050	0.070
CONTINLL	0.480	0.214	0.000	0.051	0.095	0.160
SELCON3	0.460	0.156	0.000	0.006	0.129	0.249

1 is regular, 2 is distorted secondary structure

The programs CDSSTR, CONTINLL and SELCON3, accessed using DICHROWEB predicted that the overall secondary structure of NTC was 62-79% α-helix, 0.6-14% β-strand or β-turn and 7-25% unordered structure.

CDSSTR (Johnson, 1999) uses the variable selection method by performing all possible calulations using a fixed number of proteins from the reference set. The program identifies proteins that contain characteristics that are not reflected by the test protein and removes them from the basis set. The singular value decomposition (SVD) algorithm assigns secondary structure.

CONTINLL (Provencher and Glockner, 1981) intitially used ridge regression to fit a linear combination of spectra of known composition to match a spectrum of unknown composition. However, the latest version uses the locally linearised model (van Stokkum *et al.*, 1990) where basis set proteins are screened and those that contain the lowest RMS deviation from the spectrum of unknown composition are selected. This is to ensure that only spectra which closely match the unknown spectrum are used.

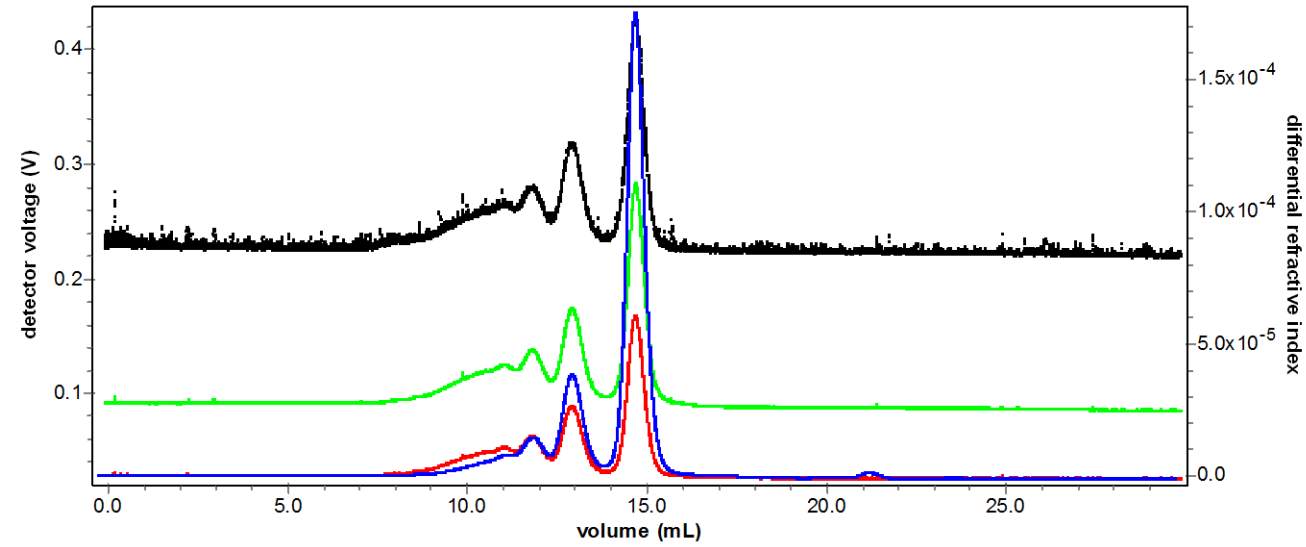
In order to predict secondary structure SELCON (Sreerama and Woody, 1993) incorporates the method in which successive iterations achieve a root mean square (RMS) noise of 0.0025. In the first step, the algorithm assigns an initial guess at the composition using the SVD method of Hennessey and Johnson (1981). In the second step, the initial starting compositions are repeated using the SVD calculation until a merged result is obtained. The results are then constrained to 1 and no composition less than -0.05. The results are then further constrained by the helix limit theorem.

Sreerama and Woody (2004) found that CDSSTR performs better than the other two programs in determining correct secondary structure. The secondary structure analysis results for CDSSTR were 79% α-helix, 14% β-strand or β-turn and 7% unordered structure. Together these are in excellent agreement with the proposal the NTC is predominantly a structured α-helical protein.

4.3.7 Investigation of the oligomeric state of NTC using gel filtration chromotography

The oligomeric state of NTC was determined using gel filtration chromatography, coupled with light scattering analysis. As discussed earlier, in section 4.3.5, the NMR spectra of NTC (figures 4.6 and 4.7) were indicative of a self-associating protein. To test this, size-exclusion chromatography was performed using a Superdex^{TM} 200 10/300GL column connected to a UV detector, a miniDAWN TREOS, multi-angle laser light-scattering detector and an Optilab rEX refractive index detector. As light scatters from the moving macromolecules, this motion relays randomness to the phase of the scattered light. When the scattered light from two or more particles is added together, there is a changing interference. This results in time-dependent fluctuations in the intensity of the scattered light (measured in volts). The intensity of the scattered light is proportional to the concentration of the macromolecules in solution. The time-dependent fluctuations in the scattered light are directly related to the rate of diffusion of the molecule through the solvent. The fluctuations are then analysed to determine a hydrodynamic radius for the sample from which the molecular mass of the macromolecule can be calculated (Wyatt, 1993). Bovine serum albumin or BSA (figure 4.11A) was used as the reference protein and the NTC data (figure 4.11B) was normalised against the reference. The molecular mass of NTC was calculated using the ASTRA V 5.3.2.21 software and it was predicted to be 34470 Da (NTC monomer is 18341 Da), which suggested that NTC is almost certainly a dimer in solution at 1 mg/mL. Therefore, the line broadening that was evident in the NMR spectra was most likely due to the self-association of NTC.

A



B

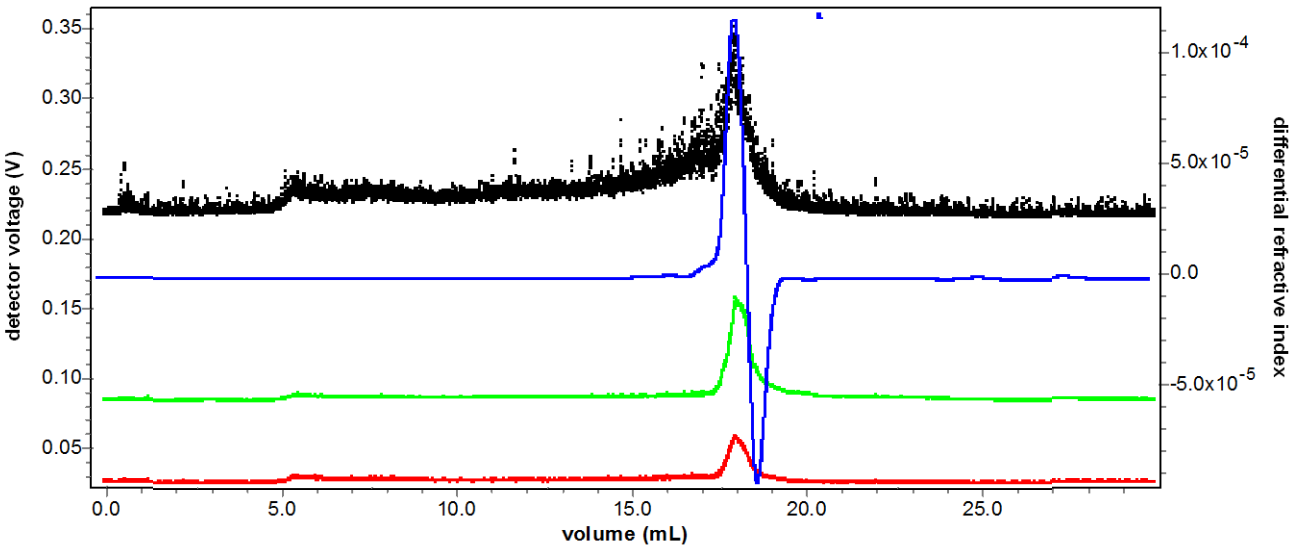


Figure 4.11 Multiple angle light scattering analysis of **A** Bovine serum albumin (BSA) and **B** HCS NTC using gel filtration chromatography and the ASTRA V software. The detectors used in ASTRA were detector 1 –, 2 –, 3 – and the refrective index –. A Superdex^{TM} 200 10/300GL column connected to a UV detector, a miniDAWN TREOS, multi-angle laser light-scattering detector and an Optilab rEX refractive index detector were used.

4.3.8 *De novo* predictions for the structures of NTC or 159-314 N-HCS and 1-151 N-HCS

In Chapter 3 it was shown that 159-314 N-HCS or NTC interacts with C-HCS. It was shown that 1-151 N-HCS was either unstable or not expressed in yeast as confirmed by Western blot analysis. The Phyre program uses a library of known protein structures obtained from the Structural Classification of Proteins (SCOP) database and the structures deposited in the Protein Data Bank (PDB). The amino acid sequence of the protein of interest is submitted and scanned against the nonredundant sequence database, and a profile is constructed (Kelley and Sternberg, 2009). Phyre was used to predict the structure of NTC as a central beta sheet surrounded on each side by alpha helices (figure 4.12) and this correlated with the secondary structure composition determined by the CD analysis of NTC. In contrast, the Phyre prediction for 1-151 N-HCS was composed of random coil (data not shown).

Phyre is an automated 3D model builder, which uses profile-profile matching and secondary structure. The models generated by Phyre rely on homology modelling. Phyre provides an estimated precision value for each resulting structure and also provides confidence values of the model. Estimated precision is based on the Phyre server benchmark on what is a true homolog. For example an estimated precision score of 95% indicates that 95% of sequences that received this score or higher were true homologs according to the SCOP database (Kelley and Sternberg, 2009). The confidence values are calculated by the Phyre server by creating a weighted average of the information from evolutionary trace, sequence conservation, and known template functional sites for each position in the query sequence. These positions are normalised between 0 and 9, with 9 being the highest confidence prediction of a functionally relevant residue (Kelley and Sternberg, 2009). Phyre was used to predict the structure of NTC, G159-R314 N-HCS and M1-E151 N-HCS. The NTC structure was aligned with Class I glutamine

amidotransferases, a flavodoxin-like protein, with an estimated precision of 95 %, meaning that the precision of the Phyre model was high. This was despite the low 9 % identity with the matched protein. The confidence in the model was rated as high. Phyre predicted NTC to be structured and correlates with the CD data. In contrast 1-151 N-HCS was predicted to be unstructured as mostly it contained random coil, a finding that is in agreement with the Lee *et al.* (2009) study that used the CDNN program and their CD data. This may explain the degradation or lack of expression of this protein observed in the yeast system.

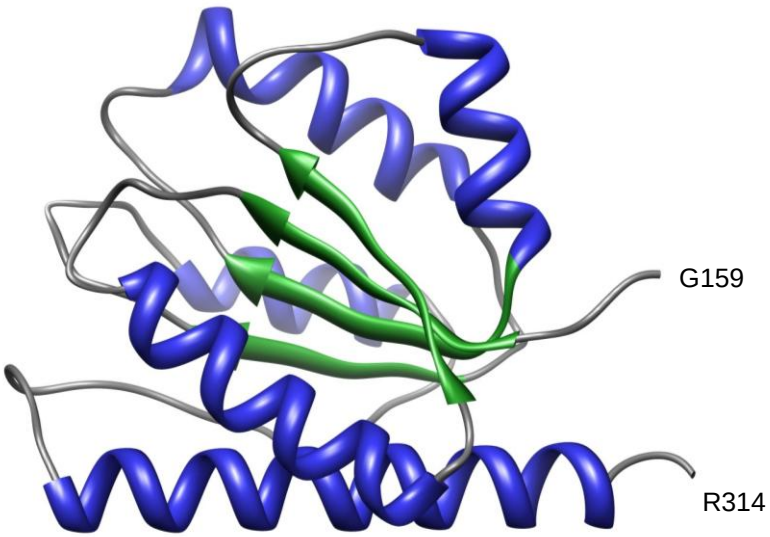


Figure 4.12 Tertiary protein structure prediction of NTC (159-314 N-HCS) using Phyre version 0.2. β strands are coloured in green and α helices in blue Figure was generated using UCSF Chimera

A qualitative test to determine the integrity of the structural model by studying the conformation of the predicted protein is by using the Ramachandran plot (Ramachandran *et al.*, 1963). The Ramachandran plot method studies angles ψ (Phi) against φ (Psi) of amino acid residues in protein structure in order to determine its plausibility. Ramachandran plot analysis indicates whether the structure generated by the modelling

program has used realistic polypeptide conformations and MOLEMAN2 (http://xray.bmc.uu.se/cgi-bin/gerard/rama_server.pl) was used in this study to determine that. The drawback of using RMSD (root mean square deviation) alone is that a good model can have a high RMSD value due to regions in the structures that do not align. RMSD values alone do not verify whether a generated model is a good model. Ramachandran plot (Ramachandran *et al.*, 1963), which analyses the conformation of the predicted protein, was used to determine whether the NTC model is a plausible structure. Like RMSD values, it does not verify whether one has a good model but it can determine whether the produced 3D structure is plausible. The Ramachandran plot method studies angles ψ (Phi) against φ (Psi) of amino acid residues in protein structure in order to determine its plausibility. The MOLEMAN2 (Kleywegt and Jones, 1996) program employs the Ramachandran plot and using this program it was found that the Phyre NTC model contained 2.3 % outliers. An average 2.0 Å structure contains 0-5 % outliers (Kleywegt and Jones, 1996) and therefore the Phyre derived NTC model is a plausible 3D structure.

Another alternative that may increase the confidence in the model is using an independent method of modelling such as Fugue (http://tardis.nibio.go.jp/fugue) or LOOPP (http://cbsupapps.tc.cornell.edu/loopp.aspx) to obtain a proposed structure of NTC. A result showing a similar structural prediction would increase the confidence in the structure presented here.

Although the Phyre model was an encouraging starting point, comparing the predicted structure with models of NTC from programs such as Fugue and LOOPP was performed in order to verify the model. The Fugue NTC model contained residues G159-S287 and the LOOPP NTC structure contained residues G159-L217. Only the Phyre model contained all the residues that encompass NTC (G159-R314) even though the input sequences were the same. Table 4.3 summarises the comparison of different sized fragments of the Phyre NTC model compared to the Fugue and LOOPP NTC models. The fragment boundaries

coincided with secondary structure boundries. Analysis using the MatchMaker feauture in Chimera showed that the structures are more structurally similar in the N-terminal end of the models. The generated Phyre NTC model was found to be structurally similar to the Fugue and LOOPP models in that it was predicted to be composed of a central β sheet flanked by α helices. The β strands however, did not align in the same direction as the LOOPP model, which contained an anti parallel strand that the other models did not.

Table 4.3 Comparison of the fragments of the Phyre NTC model with Fugue and LOOPP NTC models

Modelling Program	NTC fragment	No. of beta strands	No. of alpha helices	RMSD* value (Å)
Phyre matched with Fugue	163-246	4	3	7.496
	163-237	3	3	7.703
	163-219	3	2	6.573
	163-213	2	2	6.476
	163-197	2	1	5.403
Phyre matched with LOOPP	163-237	3(4 LOOPP)	3	9.575
	163-219	3	2	6.941
	163-213	2	2	5.024
	163-197	2	1	4.110

*RMSD is root mean square deviation

Figure 4.13 shows the top structures produced by each of the three modelling programs and the overlay of their structures using the MatchMaker function in UCSF Chimera. The NTC models show a similarity in structure in that there is a central β sheet flanked by α helices. Because the Phyre model contained all the residues that encompass NTC and the other models did not, the major difference in the structures is the presence of a long C-terminal α helix. The central β sheet in both the Phyre and Fugue models contained parallel β strands and in the LOOPP model contained two parallel and one anti-parallel β strand. The structural similarities in the models provide added confidence in the model.

The data from the secondary structure analysis of NTC CD spectrum using CDSSTR, CONTINLL and SELCON3 corresponds with the data generated from the structural models of NTC in that it is composed mostly of α-helices. The minor structural composition is β-strands and there is a small portion of unordered regions.

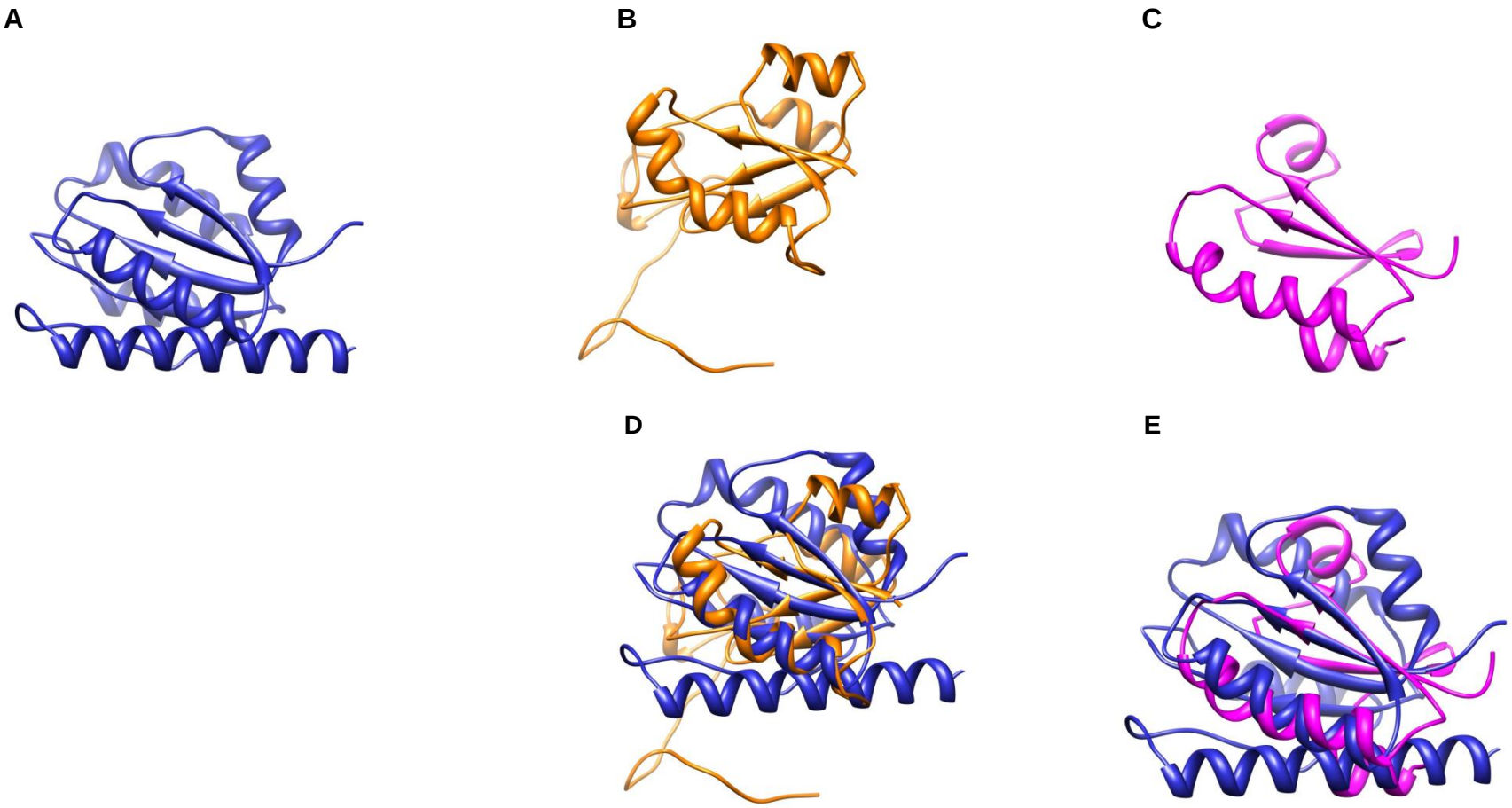


Figure 4.13 Tertiary protein structure prediction of NTC (159-314 N-HCS) using **(A)** Phyre version 0.2 (blue), **(B)** Fugue version 2.0 (orange) and **(C)** LOOPP version 4.0 (pink). **(D)** is the overlay of A and B and **(E)** is the overlay A and C using the MatchMaker function in UCSF Chimera. Figure was generated using UCSF Chimera

The Fugue program utilises sequence-structure comparison by recognising distant homologues. Fugue scores amino acid matching and insertions/deletions dependent on the local environment of each amino acid residue in a known structure. A sequence of the protein is submitted and Fugue scans a database of structural profiles, calculates the sequence-structure compatibility scores followed by the production of a list of potential homologues and alignments (Shi *et al.*, 2001). Fugue was used to predict the structure of NTC and the Z-score normalised by sequence divergence was 9.93. With Fugue a Z-score above 6 signifies 99% confidence.

LOOPP utilises a group of candidates for templates, builds an atomically detailed model from each of the templates and then ranks the models. The progam impliments a hierarchical selection process where the best model is selected using a diverse set of signals. After a screening of template candidates from the protein data bank LOOPP fine-tunes the selection to a few models. The proposed structures are then merged to give a global fitness measure using linear programming (Vallat *et al.*, 2009). LOOPP was used to predict the structure of NTC. The resulting predicted structures were ranked according to data obtained from the most confident tree to the least confident tree. LOOPP uses several decision trees for identifying and scoring final models. The structure that was used in this study was from the most confident tree. NTC structure was aligned with one of the chains of a response regulator receiver protein from *Pseudoalteromonas atlantica*.

In order to further verify the NTC Phyre structural model the residues found mutated in MCD (L216, L237) and the novel residues (L166, L206, W210) isolated in Chapter 3 that when mutated disrupted the interaction of HCS and hPC107 were mapped onto the model (figure 4.14).

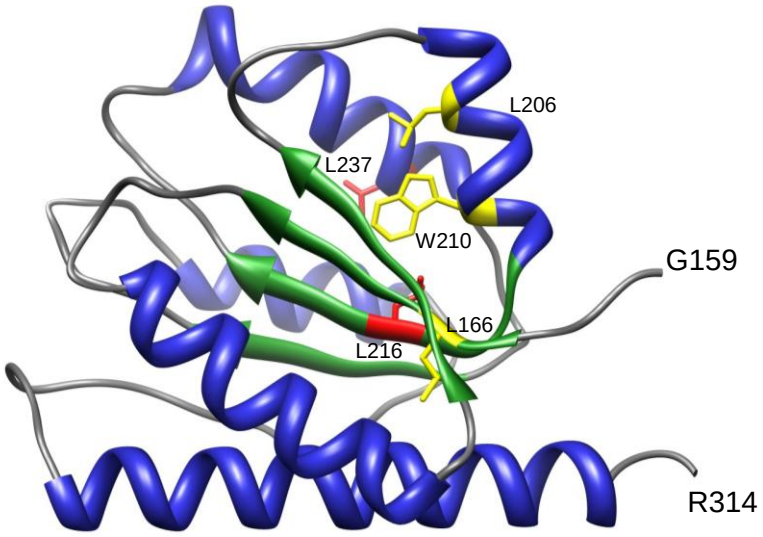


Figure 4.14 Tertiary protein structure prediction of NTC (159-314 N-HCS) using Phyre version 0.2. Showing residues that when mutated give rise to MCD (red) and the novel residues of HCS (yellow) isolated in Chapter 3 that disrupt the interaction of HCS and its protein substrate
Figure was generated using UCSF Chimera

All of the mutants were located in close proximity to each other on the structure. The mutations L216R and L166R are located on the same β-strand and L206P and W210R are located on the same α-helix. All of the mutations could all be contributing to the hydrophobic core and thereby may assist in stablilising the structure of the protein.

4.4 Summary and future work

There are currently no high resolution structures of HCS or its domains. HCS has a long N-terminal extension, which is not present in class I and II BPLs and its structural elucidation is important as this study has shown that it has a role in protein substrate binding. In Chapter 3, it was shown that the MCD mutations that are located in the NTC disrupted the interaction between HCS and its substrate hPC107. This highlights the

functional importance of the N-terminal extension in HCS, which is absent in other BPLs. Due to the important role of NTC and the lack of structural information on HCS, in this Chapter, the elucidation of its structure was attempted using NMR.

For the expression and purification of NTC different NTC constructs were transformed into their respective bacterial or yeast host strains. Optimisation of the expression of the fusion proteins was performed, however only the GST-fused constructs expressed in *E. coli* BL21 were successfully expressed. Of these only the GST-NTC (159-314) was soluble. Once the appropriate fusion protein was selected the next issue was that low amounts of protein were purified; only 2.1-2.8 mg of NTC was purified from a 2 L culture. To overcome this, optimisation of the expression of GST-NTC was performed by varying the expression temperatures, incubation times and IPTG concentrations. This did not improve yield. Due to this, four purifications of NTC were pooled for each NMR experiment. After the purification, NTC was stored at 4°C for 18 hours until the NMR experiments were conducted. The NMR experiments were conducted at 25°C and these ran for 12 to 24 hours. After these experiments it was determined that NTC precipitated out of solution and since heteronuclear NMR spectra required for structure determination can take up to 3 days to record and multiple such spectra are required, it is important that the sample be stable under these conditions for 4 to 5 days at 25°C. The lack of many cross peaks in the NMR spectra may be due the self-association of NTC.

The CD spectrum of NTC exhibited a spectrum typical of a structured α-helical protein and this was confirmed by the analysis of the spectrum using CDSSTR, CONTINLL and SELCON3 programs. Performing CD at 90°C in order to denature the protein showed that the folded NTC unfolded or melted at this temperature. The CD spectrum at 20°C after melting showed that NTC did not completely re-fold. A melting experiment was performed using CD by monitoring the ellipticity at 220nm at every 1°C ranging from 20°C to 90°C and

the melting temperature (T_m) was determined to be 55.7°C. This observation further validated that NTC was folded as heating the sample induced unfolding of the protein and the temperature at which half of the protein is unfolded (the melting temperature) could be determined.

The oligomeric state of NTC was determined using gel filtration chromatpgraphy as the NMR data were indicative that NTC may self-associate. Size-exclusion chromatography confirmed this observation as NTC was determined to likely be a dimer in solution as the molecular mass of the NTC was double that of monomeric NTC. Therefore, the evident line broadening that was observed in the NMR spectra was due to the self-association of NTC. Due to this observation, transverse relaxation-optimised spectroscopy (TROSY) would be more suitable in the structural analysis NTC.

The problems that limit convensional NMR studies of proteins larger than 25 kDa are signal overlap, which is due to the large number of resonances and NMR signals of larger molecules relax faster and this leads to line broadening and decreased spectral sensitivity (Fernandez and Wider, 2003). In general, the overlap of signals in the NMR spectra can be surpassed by decreasing the number of resonance lines by selecting appropriate isotope-labels, such as ^{2}H, ^{13}C and ^{15}N (Gardner and Kay, 1998; Otomo *et al.*, 1999; Fernandez and Wider, 2003). However, the limitation caused by transverse relaxation presents a greater technical challenge. Relaxation is mainly due to presence of hydrogen atoms and their replacement by deuterons in turn reduces transverse relaxation significantly, which results in increased resolution and significant sensitivity gains. However, protons contribute significantly to the structure and produce the most sensitive NMR signal and therefore measuring totally deuterated proteins is not ideal. Therefore C-H groups of proteins are usually partially deuterated and dissolved in H_2O solution to replace deuterons with protons in exchangeable sites, where each amino acid residue is protonated at the backbone amide position (Fernandez and Wider, 2003). However, solely

using deuteration, does not completely extend the application of solution NMR for larger proteins or proteins were line broadening is observed. Only the introduction of TROSY (Pervushin *et al.*, 1997) reduces relaxation to achieve satisfactory line widths and sensitivity in such NMR experiments. TROSY works best with deuterated proteins and is particularly suited for applications to protonated amide groups. With the use of TROSY the effective relaxation of the measured signal during the pulse sequence and during data acquisition can be reduced. This enables the measurement of high-quality spectra that otherwise can not be produced with conventional NMR studies (Fernandez and Wider, 2003). While NMR experiments using TROSY would be more suitable in analysing NTC this however was beyond the scope of this thesis.

Since the structure of NTC was not obtained using NMR, structural models of NTC were produced using the programs Phyre, Fugue and LOOPP and the predicted the structure contained a central beta sheet flanked by alpha helices. The NTC model secondary structure had mostly a helical content which was in agreement with the CD spectrum analysis using CDSSTR, CONTINLL and SELCON3. Although this model is a good starting point for the structural analysis of NTC, a high resolution structure of NTC is required and as I have presented, TROSY experiments would be appropriate in achieving this or alternatively X-ray crystallography.

Although structural information of the N-terminal region HCS could not be elucidated using NMR, functional studies of the region were carried out. Chapter 5 characterises the role of the NTC in protein substrate binding using surface plasmon resonance (SPR).

CHAPTER 5

The Characterisation of MCD HCS mutants

5.1 Introduction

In Chapter 3 it was shown that the interaction between N-terminal and C-terminal regions of HCS were not disrupted by the MCD mutants L216R and L237P HCS using the Yeast Two-hybrid assay. It was also shown that the interaction between MCD mutant HCS and its substrate hPC107 was disrupted. From these findings, I concluded that the mechanism by which most of the HCS mutants affect catalytic activity is by disruption of the second partial reaction, where the interaction of HCS with its protein substrates occurs. I subsequently decided that it was important to confirm these genetic studies with biophysical analysis using purified protein.

Therefore the aim of this Chapter was to characterise the interaction between HCS and its substrate hPC107 using Surface Plasmon Resonance (SPR).

5.2 Specific Methods

For the investigation of the effect MCD mutants have on the interaction between HCS and the hPC107 using SPR required purified material for accurate analysis.

5.2.1 Expression and purification of hPC107

E. coli BL21 containing pGEX-hPC107 in LB (0.1 mg/mL Ampicillin) culture was prepared and incubated at 30°C until the culture reached an OD_{600} of 0.6. Induction of GST-hPC107 was performed by adding 0.1 mM IPTG and the culture incubated a further 2 hours. Cells were harvested and resuspended in 20 ml of lysis buffer (Tris buffered saline, pH 8.5) per 1 L culture and lysed with a cell disruptor. Insoluble material was removed by centrifugation at 21 000 x g for 20 min. The supernatant was filtered using a 0.8 µm followed by a 0.45 µm filter. Purification of hPC107 was performed as described in Chapter 2 section 2. Thrombin cleavage was performed at 37°C for 2 hours followed by the elution of cleaved hPC107 using TBS, pH 8.5.

5.2.2 Expression and purification of wild type HCS, L216R HCS and L237P HCS

A hexahistidine tag engineered onto the C-terminus of the recombinant enzymes facilitated protein purification using immobilised metal affinity chromatography. The pET vectors for the expression of wt HCS-His6, L216R HCS-His6 or L237P HCS-His6 were transformed in *E. coli* strain Rosetta (DE3)pLysS^{RARE} (Novagen). Rosetta(DE3)pLysS^{RARE} strains enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. This strain contains the pLysS plasmid, which produces T7 lysozyme, an inhibitor of T7 RNA polymerase, by destabilising the open complex during initiation (Stano and Patel, 2004) and this reduces basal level expression of proteins that are toxic in *E. coli*. Cells were grown to OD_{600} 0.6 in 2x YT (0.1 mg/mL Ampicillin) at 30°C and expression was performed using 0.1 mM IPTG at 23°C or room temperature for a further 18 hours. Cells were

harvested and resuspended in 20 ml of lysis buffer (300 mM KCl, 50 mM KH_2PO_4, 5 mM imidazole, pH 8) per 1 L culture and lysed with a cell disruptor (Microfluidics). Insoluble material was removed by centrifugation at 25 000 x g for 20 min using a Beckman centrifuge. The supernatant was then filtered using a 0.8 µm followed by a 0.45 µm pore-sized filter.

Purification of wild type and mutant HCS was performed using the Profinia™ Protein Purification System (BIO-RAD) with a Bio-Scale™ Mini Profinity™ IMAC (immobilised metal ion chromatography) 5 ml cartridge. The purification was carried out as specified by the user manual. The buffers employed were those indicated in the manual except the Native IMAC wash buffer 2 contained 40 mM imidazole instead of 10 mM. The desalting or storage buffer used contained 50 mM Tris, pH 7, 1 mM EDTA and 5 % (w/v) glycerol. SDS-PAGE analysis was performed to analyse the purification. The ^3H biotin incorporation assay was performed to determine enzyme activity.

5.2.3 Circular Dichroism (CD) Spectroscopy

Protein samples were buffer exchanged by dialysis into 10 mM sodium phosphate buffer, pH 8.0, by dialysis using regenerated cellulose tubular membrane (Fisher biotech, Australia). Following centrifugation at 17 900 x g for 1 minute the concentrations of the samples were determined using the Bradford assay. Wild type, L216R and L237P HCS at a concentration of 0.2 mg/mL were analysed using a Jasco J-185 CD spectrophotometer. CD was carried out as in Chapter 2, section 2.2.6. Each spectrum was recorded with an average of 5 scans with subtraction of the buffer blank 10 mM sodium phosphate buffer, pH 8.0.

5.2.4 Surface Plasmon Resonance (SPR)

Surface plasmon resonance was performed using a Biacore^{TM} T100 instrument (GE Healthcare). HCSs were initially immobilised on a CM5 sensor chip using amine-coupling chemistry by activating the sensor surface by injecting N-ethyl-N`-(3-diethylamino-propyl)carbomiide and N-hydroxysuccinimide. Following which 0.3 mg/ml of each HCS in 10 mM sodium acetate buffer (pH4.6) was injected onto the activated surface and 1 M ethanolamine hydrochloride was injected. Approximately 6500-7000 resonance units (RU, 5 ng/mm^2) of each HCS were immobilised onto the sensor surface.

Each binding experiment was performed at a flow rate of 30 µl/min at 25°C with HBS running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl) containing 3 mM EDTA for biotin binding or no EDTA for biotin and MgATP binding. Various concentrations of biotin (0-500 µM) were prepared in the running buffer and injected across the sensor chip. The binding of MgATP was determined by injecting 125 µM MgATP over the surface of the chip. Presumed biotinyl-5´-AMP synthetase activity was determined by injection of biotin:MgATP (125:125 µM) in running buffer. Once biotinyl-5´-AMP was formed and therefore holo-HCS was formed, varying concentrations of apo-hPC107 (0.039 - 40 µM, diluted in running buffer) were injected onto the HCS-immobilised chip. A 1:1 binding model was used to calculate the association rate constant (k_a), dissociation rate constant (k_d) and equilibrium dissociation rate constant (K_D).

5.3 Results

A variety of protein based techniques were attempted to complement the genetic data described in Chapter 3. These are shown in a flow chart in figure 5.1. For the GST pull-down approach, GST fused N-HCS, 1-151 N-HCS and 159-314 N-HCS were successfully cloned and purified. However, the C-terminal region of HCS could not be recombinantly produced in either *E. coli* or yeast. Varying growth temperatures and IPTG concentrations were investigated. The conditions included lowering the expression temperature from 37°C to 30°C and to 25°C. The IPTG concentrations that were used included 0.2, 0.1, 0.05, 0.025 and 0.01 mM. Due to this technical challenge, the GST-pull down experiment could not be completed. An alternative approach that was attempted involved engineering a TEV site (figure 5.1) into HCS to allow digestion of HCS into two halves. The products could then be resolved by gel filtration to determine whether they co-elute under native conditions. This approach would supersede the requirement to produce individual fragments. Site directed mutagenesis was used to change A328N, V329L, L330Y, C331F, V333G. With the unchanged residues E327 and Q332 the consensus sequence for the TEV site ENLYFQG was introduced. After mutagenesis HCS[TEV]-His6 was expressed and purified using a nickel column (specific activity was 0.5 nM holo PC formed/min/mg, which was three times lower than wild type HCS). Various enzyme:substrate ratios were trialled, however, after incubation with TEV protease the N-terminal region of HCS was degraded by non-specific protease activity. Due to this degradation the gel filtration step could not be completed. In this Chapter the technique that was used to determine the effect of the MCD mutants on the interaction between HCS and hPC107 was SPR.

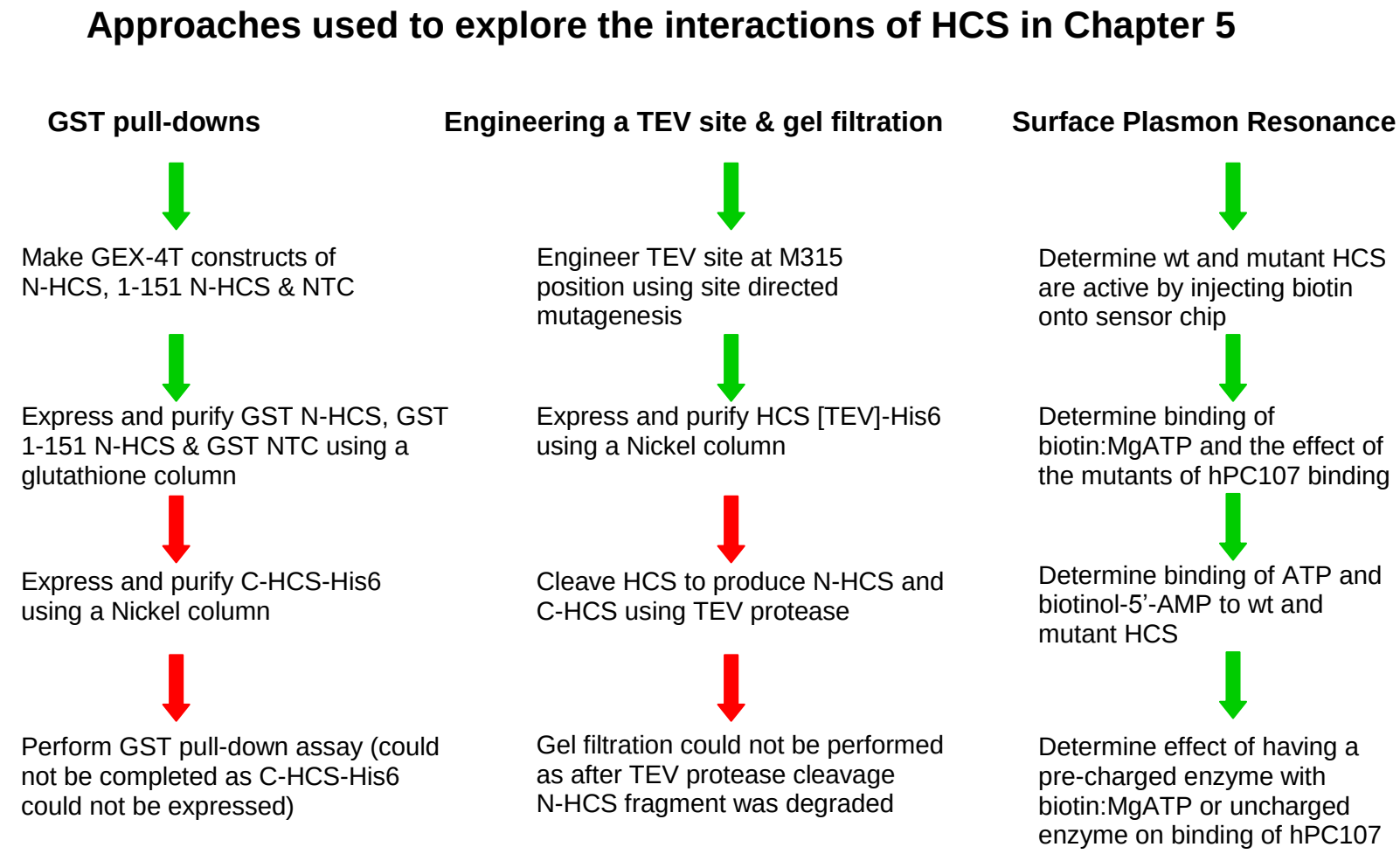


Figure 5.1 Flow chart summarising the approaches for exploring the interactions of HCS

5.3.1 Expression and purification of hPC107

In order to determine the effect of the MCD mutations on the interaction between HCS and its substrate hPC107 using SPR, purified hPC107 was required. *E. coli* BL21 containing pGEX-hPC107 was used to express GST-hPC107. Purification of hPC107 was performed using the Profinia™ Protein Purification System (BIO-RAD) using the Bio-Scale™ Mini Profinity™ GST 5 ml cartridge. The purification method including on column cleavage with thrombin is outlined in section 5.2.1. Fractions from the purification were analysed using SDS-PAGE (figure 5.2) and lane PF shows purity of hPC107. Protein concentration was calculated using the Bradford assay and 0.95 mg/mL or 4.7 mg total hPC107 from 2 L culture was purified using this method.

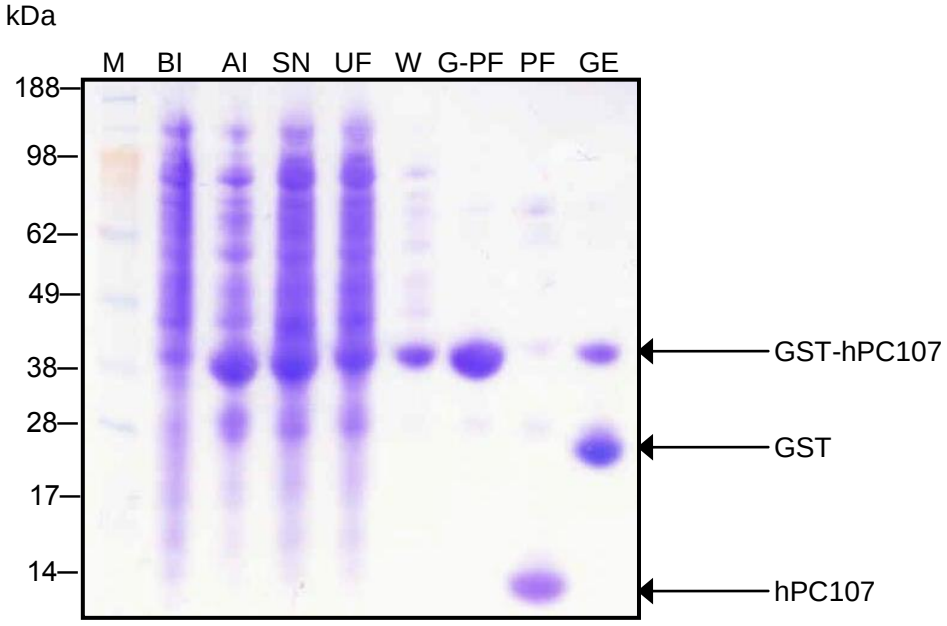


Figure 5.2 Purification of hPC107 using the Profinia™ Bio-Scale™ Mini Profinity™ GST cartridge. M - Molecular Weight marker (SeeBlue pre-stained marker, Invitrogen), BI - before induction, AI - after induction, SN - supernatant, UF - unbound fraction, W - wash, G-PF - GST fused purified fraction, PF - purified fraction and GE - Glutathione elution

5.3.2 Expression and purification of wild type HCS, L216R HCS and L237P HCS

Wild type HCS was initially expressed in LB at 37°C, however due to a lack of observed expression (confirmed by SDS-PAGE), optimisation of its expression was performed. The temperature was reduced to 30°C, however, no expression was observed even with 0.5 mM IPTG. The media was subsequently changed to the nutrient rich 2x YT and SDS-PAGE analysis revealed protein induction but enzyme assays showed the protein was inactive. To overcome this a slower expression was employed where cells were grown to OD_{600} 0.6 at 30°C and expression was performed using 0.1 mM IPTG at 23°C or room temperature overnight. Wild type HCS was determined to be active under these conditions, which were then employed for all subsequent preparations.

5.3.2.1 Optimisation of HCS purification using a higher concentration of imidazole

Purification of wild type and mutant HCS was performed using the Profinia™ Protein Purification System (BIO-RAD) and using the Bio-Scale™ Mini Profinity™ IMAC 5 ml cartridge. Each purification step was analysed using SDS-PAGE. The IMAC cartridge used was a Ni-chelated column and Profinia™ was used to purify wt HCS-His6, L216R HCS-His6 and L237P HCS-His6. To limit non-specific binding of metal-binding proteins to the resin the lysis buffer and first wash buffer contained 5 mM imidazole and wash buffer 2 contained 10 mM imidazole. The HCS proteins were eluted from the resin with 250 mM imidazole. Under these conditions a non-specific bacterial protein of 64 kDa was present in each preparation. Anion-exchange chromatography was employed to purify the proteins further, however, this led to a complete loss in protein (data not included). The Profinia™ IMAC purification method was subsequently modified and optimised as an alternative to the anion-exchange chromatography purification. The stringency of the purification was improved by increasing the concentration of imidazole in wash 2 from 10 mM to initially 20 mM imidazole and then to 40 mM imidazole (figure 5.3A-C). Figure 5.3 (A) shows that at 20 mM imidazole the contaminating 64 KDa protein was still present. However

preparations of all HCS variants using 40 mM imidazole in the second wash step did not contain this contaminating material (figure 5.3A-C). Total protein yields from 4 L of culture were 3.0 mg wt HCS, 2.5 mg L216R HCS and 3.1 mg L237P HCS.

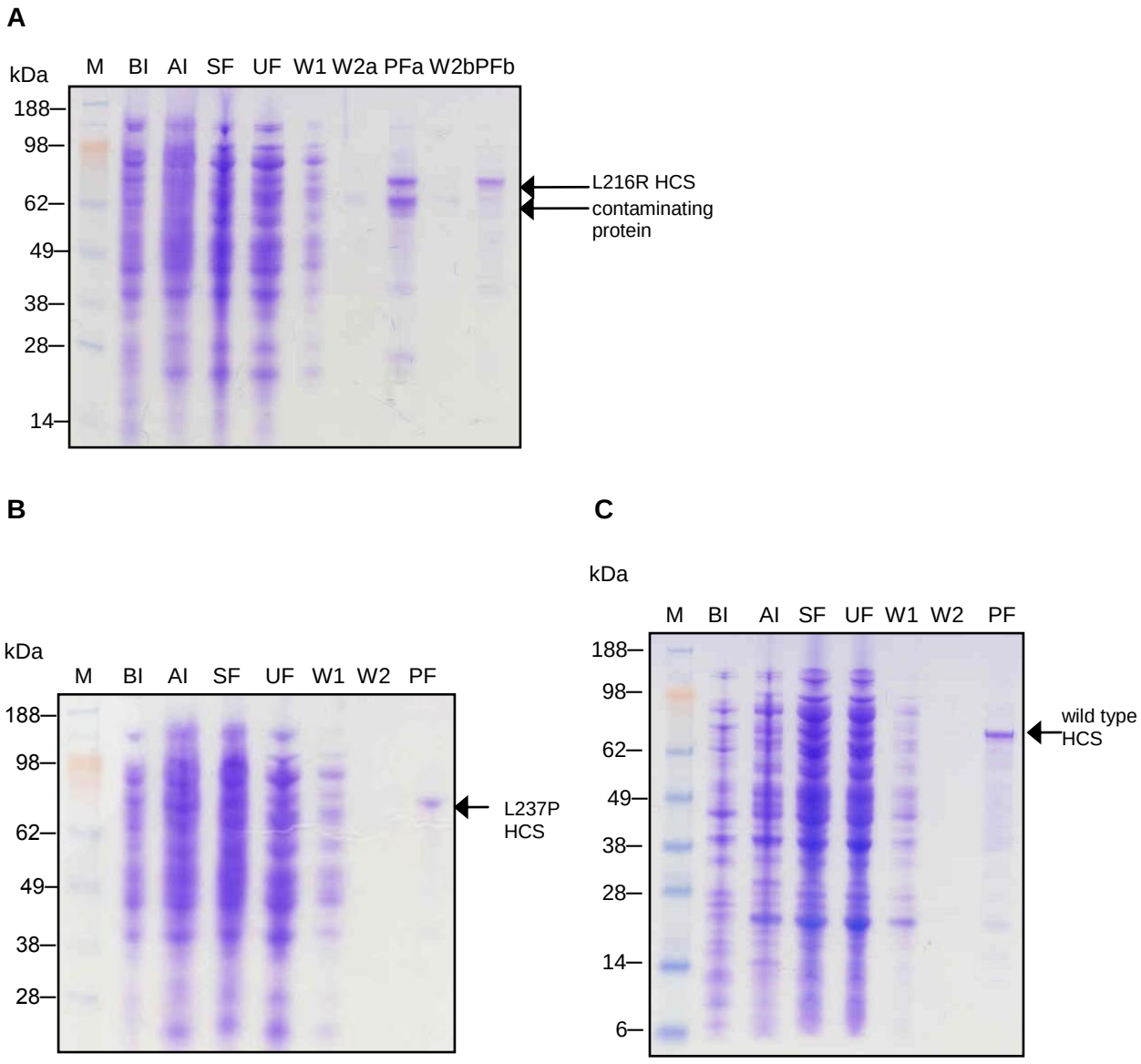


Figure 5.3 Purification of **(A)** L216R HCS-His6, **(B)** L237P HCS-His6 and **(C)** wild type HCS using the modified Profinia™ IMAC system. M - Molecular Weight marker BI - before induction, AI - after induction, SF - soluble fraction, UF - unbound fraction, W1 - wash 1, W2a - wash 2 containing 20 mM imidazole, W2b (for A) or W2 (for B and C) - wash 2 containing 40 mM imidazole, PFa - purified fraction washed with 20 mM imidazole and PFb (for A) or PF (for B and C) - purified fraction washed with 40 mM imidazole.

5.3.3 HCS activity assay

Before performing SPR, the purified wild type and mutant HCS were assayed in order to determine their specific activity (figure 5.4). Enzyme activity was determined by utilising the purified hPC107 as a substrate in an assay to measure ^{3}H biotin incorporation. Wild type HCS had a specific activity of 1.5 nmol holo hPC-107 formed/min/mg, whereas both mutant proteins contained statistically significantly compromised enzymatic activity ($p<0.01$), 3.3 % and 3.5 % of wild type for L216R and L237P, respectively. The mutants both showed above background activity and the decrease in % activities were consistent with previous studies on patient cell lines (Aoki *et al.*, 1997; Wilson *et al.*, 2005).

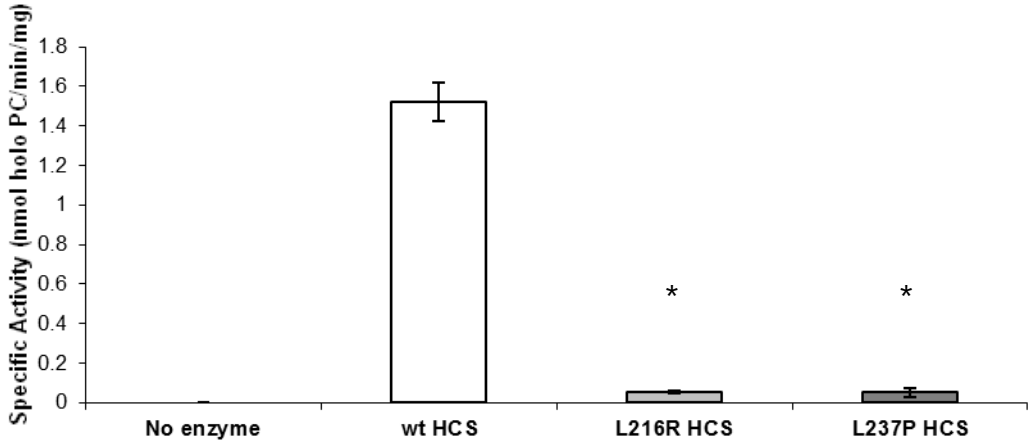


Figure 5.4 Graph showing specific activities of wild type, L216R and L237P HCS determined using the ^{3}H biotin incorporation assay. The values presented are means ± standard deviations of three samples from a single experiment.
* $p< 0.01$. The p value was calculated using the z-score.

5.3.4 Circular Dichroism (CD) Spectroscopy

In order to determine that the MCD mutant HCS did not grossly alter the secondary structure of HCS, far UV CD was performed. The HCS enzymes were dialysed against 10 mM sodium phosphate buffer, pH 8.0. Individual spectra (including 10 mM sodium phosphate buffer, pH 8.0, as a blank) were recorded as outlined in section 5.2.3. Figure 5.5 shows the overlaid CD spectra. From this analysis it appeared that the MCD mutants do not affect the gross overall fold of HCS as proteins contained overlapping spectra. The spectra were also consistent with all three proteins being structured as they did not have a spectrum typical of a disordered protein, which contains very low ellipticity above 210 nm and negative readings near 195 nm (Venyaminov *et al.*, 1993).

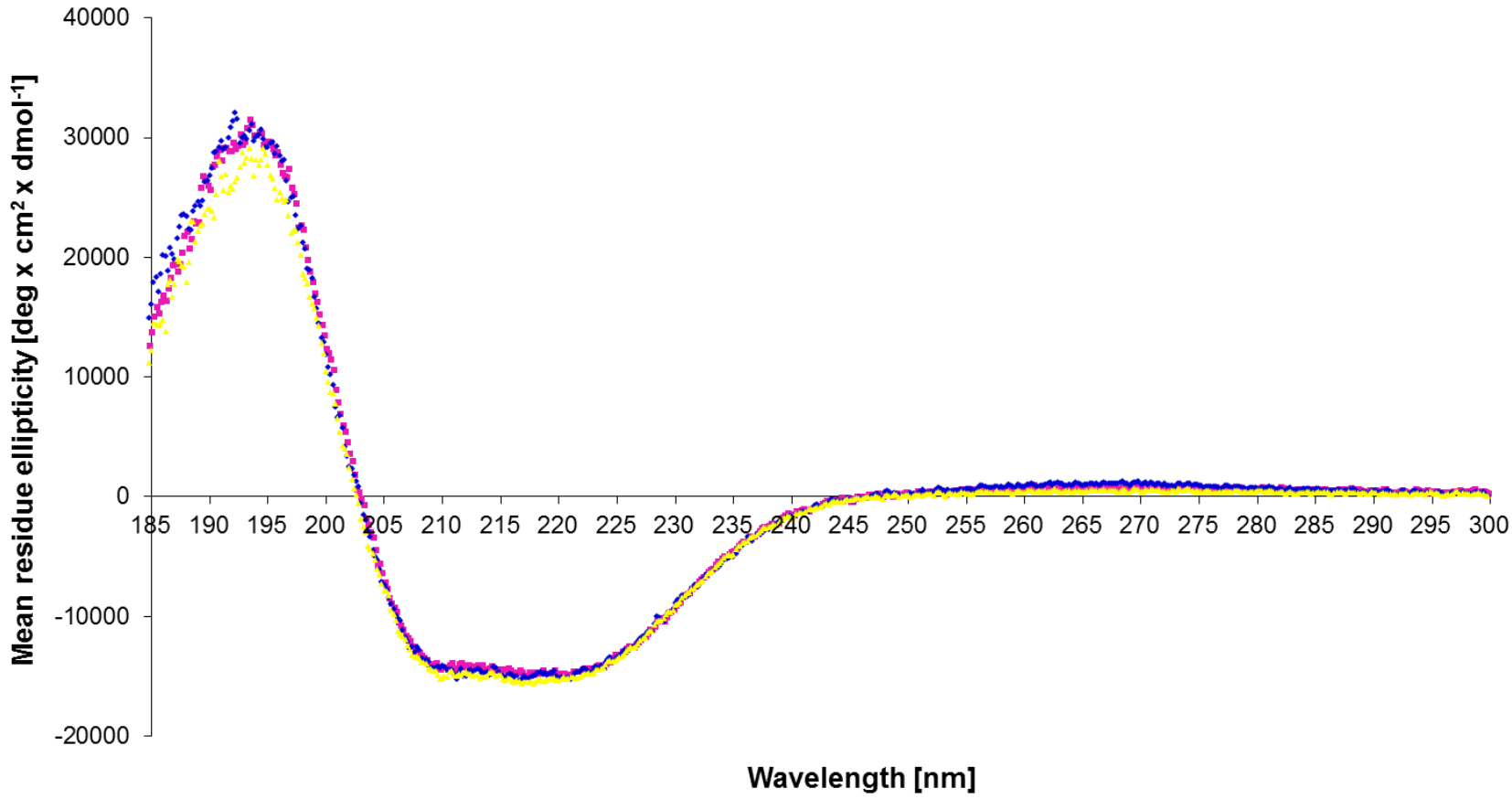


Figure 5.5 Overlaid circular dichroism spectra of wild type ■, L216R ◆ and L237P HCS ▲
Wild type and mutant HCS samples were exchanged into 10 mM sodium phosphate buffer, pH 8.0 by dialysis. The samples were diluted to 0.2 mg/mL. Individual spectra (including 10 mM sodium phosphate buffer, pH 8.0, as a blank) were recorded at 20°C across wavelengths 185 nm to 300 nm with a nitrogen flow of 7.5 L/min, scan speed of 20 nm/min and accumulation mode of 5 (average of 5 scans).

5.3.5 Surface Plasmon Resonance (SPR)

In Chapter 3, I provided genetic evidence demonstrating mutation of the NTC can effect the protein:protein interaction with a biotin domain. To further probe this result SPR (Biacore T100) was employed to monitor the molecular interactions. SPR monitors the interactions by monitoring the changes in the mass concentration of macromolecules at the surface of the sensor chip (Liedberg *et al.* 1983). In order to explore the mechanism of action of the MCD mutants and determine that the immobilised enzymes were biologically active, binding studies of ligands biotin and MgATP to wild type HCS, L216R HCS and L237P HCS were performed. Biotinyl-5′-AMP synthetase activity was determined for each protein. The molecular mechanism of the mutants was further probed by determining the association rate constants (k_a), dissociation rate constants (k_d) and equilibrium dissociation constants (K_D) of hPC107 to wild type and MCD mutant HCS.

5.3.5.1 pH scouting and immobilisation of protein

The first step performed for SPR was pH scouting where the optimal pH for immobilisation of the protein onto the sensor chip was determined. For electrostatic pre-concentration to occur, the protein must contain a net positive charge. The pH of the buffer, therefore should be below the isoelectric point (pI) of the protein. For immobilisation onto a negatively charged CM5 (coated with carboxymethylated dextran) sensor chip, the pH of the buffer is suggested to be 1 pH unit below the pI of the target protein. HCS has a predicted pI of 5.6 and, therefore, the buffer at pH 4.6 was employed for the wild type immobilisation (0.3 mg/mL). The protein was passed over the chip until 7500 RU was reached. The wild type and MCD mutant HCS were subsequently immobilised on a CM5 sensor chip using amine-coupling chemistry. Amine coupling introduces N-hydroxysuccinimide esters into the CM5 surface by modifying the carboxymethyl groups with a mixture of N-hydroxysuccinimide (NHS) and N-ethyl-N'-dimethyl-aminopropyl)-carbodiimide (EDC). The esters react with amines on the ligand and form covalent links. In each SPR experiment that was performed in this study, a blank flow cell where no protein was immobilised was present. All SPR data from each experiment was subtracted with the blank flow cell data using the Biacore Software.

5.3.5.2 Determination of the activity of immobilised HCS on CM5 chip and binding of biotin and MgATP occurs in an ordered manner

In order to confirm that the immobilised HCS had retained biological activity, biotin binding and presumed biotinyl-5´-AMP synthetase activities were measured (figure 5.6). Initially increasing concentrations of biotin were passed over immobilised wild type HCS, L216R HCS and L237P HCS (figure 5.6 (1A-1C), respectively). As figure 5.6 (1A-1C) shows, biotin bound to wild type HCS, L216R HCS and L237P HCS in a concentration dependent manner. The response was proportional to the concentration of biotin injected onto the sensor chips containing the wild type and MCD mutant HCS (figure 5.7). Comparable changes in surface plasmon resonance were observed for wild type and the two mutants suggesting equivalent affinities for biotin for the 3 proteins. The sensorgrams also demonstrated that biotin associates and dissociates from the wild type HCS and MCD mutant HCS with rapid kinetics. Due to the rapid on and off rates, steady state kinetics were required to determine the K_D. The K_D for biotin could not be determined, however, as saturation could not be reached using steady state kinetics even at the highest concentration (500 μM). The blank flow cell with no protein immobilised contained no response, which eliminated the chance of non-specific binding.

To investigate whether the immobilised proteins contained biotinyl-5´-AMP synthetase activity either MgATP (125 μM), biotin (125 μM) or biotin:MgATP (125 μM) were injected onto the sensorchip (figure 5.8). The response generated for each enzyme for MgATP was no higher than 4 response units (RU), compared to 14 RU for biotin. In contrast, a synergistic response of 88 RU was observed for biotin:MgATP. Unlike biotin, MgATP and biotin produced a species that remained on the surface of the sensor chip. This most likely represents the formation of biotinyl-5´-AMP. These data suggest that HCS binds ligands in an ordered binding reaction whereby biotin binds prior to MgATP. This is the first report analysing the reaction mechanism using SPR and this observation is in agreement with findings from other biophysical studies into the HCS mechanism (Ingaramo and Beckett, 2009).

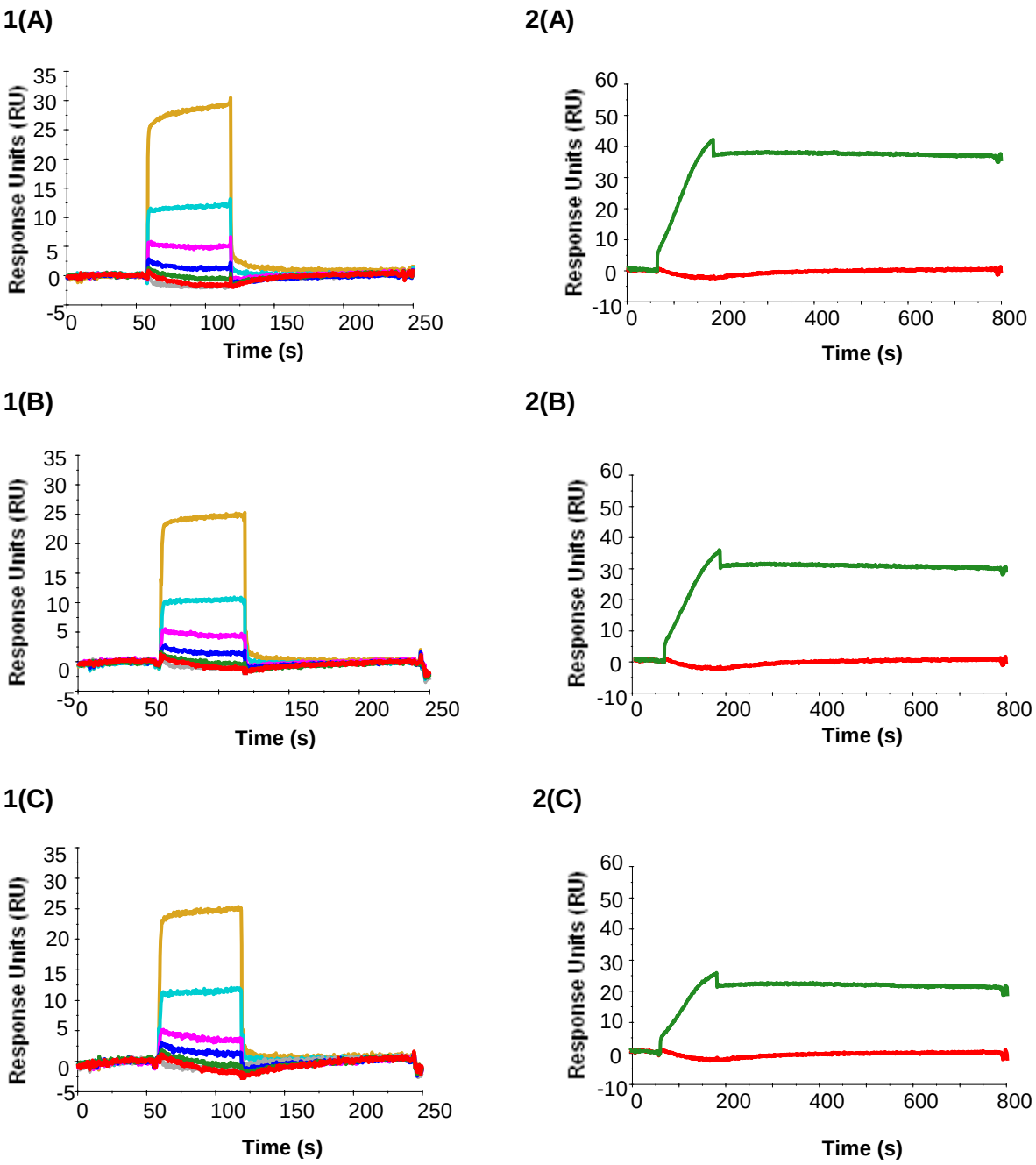
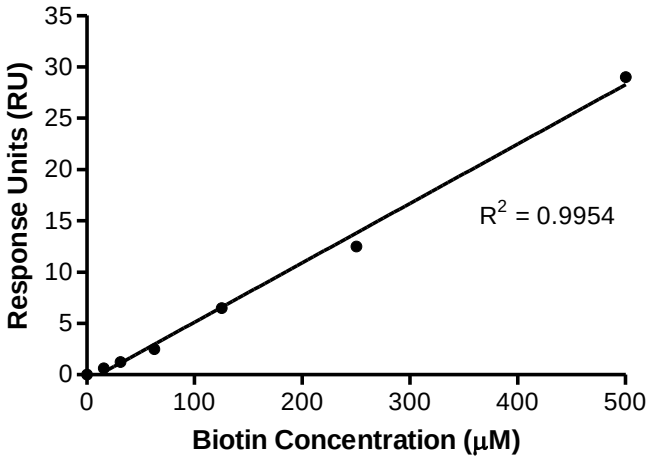
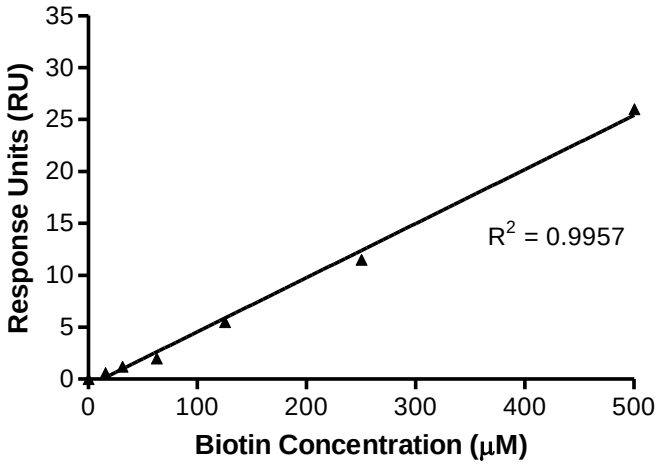


Figure 5.6 Sensorgrams showing binding of increasing concentrations of biotin (0, 15.625, 31.25, 62.5, 125, 250 and 500 µM) onto **1(A)** wt HCS, **1(B)** L216R HCS and **1(C)** L237P HCS and the measurement biotinyl-5´-AMP synthetase activity **2(A)** wt HCS, **2(B)** L216R HCS and **2(C)** L237P HCS using Biacore™. Biotin was injected after 60 seconds and the chip was washed with HBS after 120 seconds. To measure biotinyl-5´-AMP synthetase activity biotin:MgATP was injected after 60 seconds and washed with HBS after 200 seconds for a further 600 seconds. Red line represents buffer alone and green line shows response when ligands were bound. Biacore T100 Evaluation software version 1.1.1 was used.

A



B



C

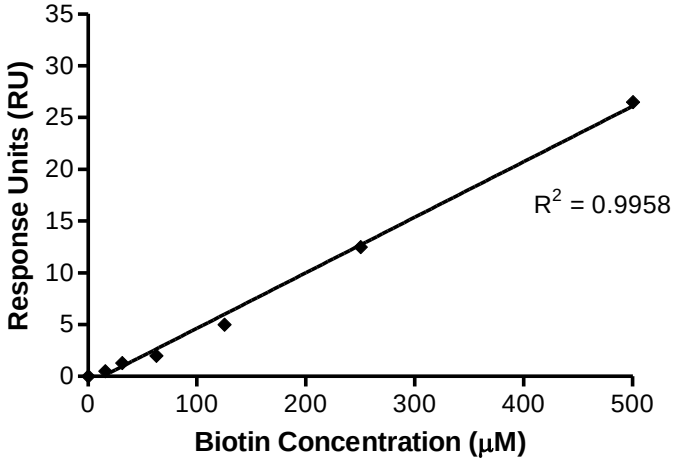


Figure 5.7 Binding of biotin to HCS occurs in a concentration dependent manner **(A)** Wild type HCS, **(B)** L216R HCS and **(C)** L237P HCS

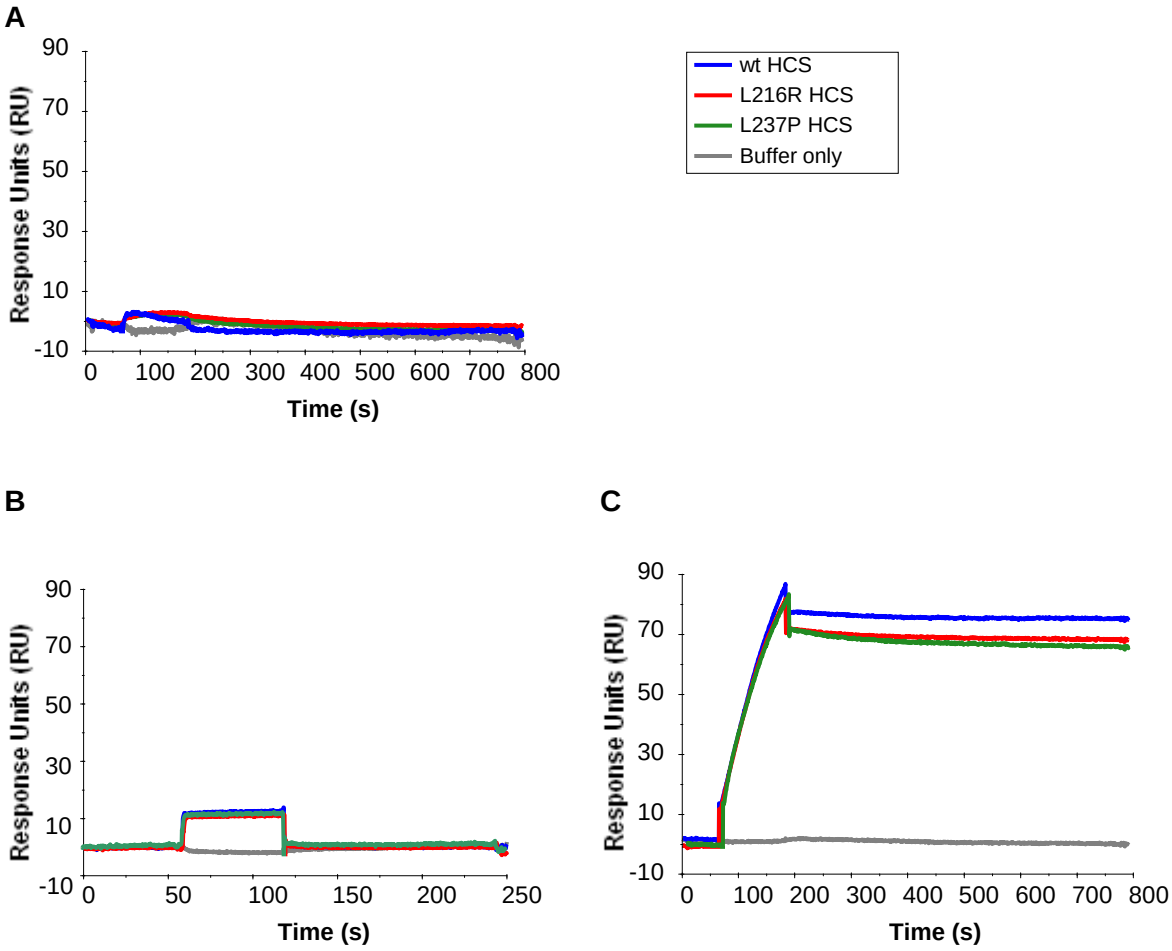


Figure 5.8 Sensorgrams of **(A)** MgATP (125 µM), **(B)** 125 µM biotin and **(C)** 125 µM biotin:MgATP injected onto a wild type and MCD mutant HCS. MgATP (A) and biotin (B) were injected at 60 seconds and washed with HBS after 120 seconds. Biotin:MgATP (C) was injected at 60 seconds and washed with HBS after 200 seconds. Biacore T100 Evaluation software version 1.1.1 was used.

I investigated biotinyl-5´-AMP synthetase activity where biotin and MgATP were both injected onto the HCS-immobilised chip (figures 5.6 and 5.8(C)) for 200 seconds before washing the chip with HBS. The formation of a slow dissociating complex was observed. This was consistent with the hypothesis that the reaction intermediate of the HCS reaction contains a slow dissociation rate from HCS. Bimolecular interaction studies using synthetic biotinyl-5´-AMP and the BPL from *E. coli* have demonstrated that the holo enzyme complex is long-lived with a half-life of 30 minutes (Naganathan and Beckett, 2007; Streaker and

Beckett, 2003). In the present study, the complex was stable during a wash for an additional 600 seconds (figure 5.6 2 (A-C)) for all 3 enzymes. In order to confidently verify that biotinyl-5´-AMP was formed on the sensor chip mass spectrometry of the formed product would need to be performed. This is technically challenging given the small amounts of material used for SPR and beyond the scope of my candidature.

5.3.5.3 Binding of biotinol-5´-AMP to wild type and mutant HCS

A control experiment was performed to address whether the compound biotinol-5´-AMP forms a slow dissociating complex with HCS. Biotinol-5´-AMP is a non-hydrolysable analogue of the reaction intermediate biotinyl-5´-AMP, and therefore could potentially be used to generate a form of holo HCS that did not proceed through catalysis. Previous work performed by Tatiana Soares da Costa (personal communication, 2010) showed that biotinol-5´-AMP formed a slowly dissociating complex with *S. aureus* BPL (*Sa*BPL), comparable to that of the biotin:MgATP complex, which also dissociates slowly. Figure 5.9 shows a representative sensorgram from three experiments where 125 µM biotinol-5´-AMP was injected over wild type and MCD mutant HCS. Biotinol-5´-AMP had similar binding kinetics to biotin with rapid association rates and dissociation rates. This finding was in direct contrast to *Sa*BPL. *In vitro* kinetic analysis showed that the compound had a 10 fold higher affinity for *Sa*BPL compared with HCS (Soares da Costa, 2010), suggesting a difference in the binding mechanism between species.

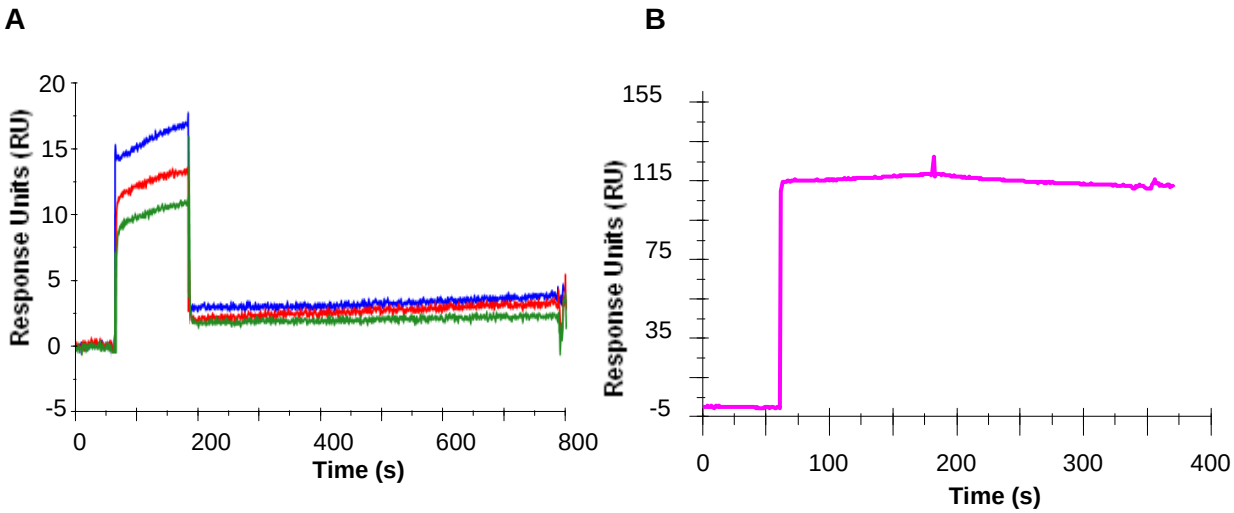


Figure 5.9 Sensorgram showing the association and dissociation of Biotinol-5´-AMP to and from **(A)** wild type, L216R and L237P HCS and **(B)** wild type *S. aureus* BPL (Soares da Costa, 2010). Biotinol-5´-AMP (125 µM for A and 50 µM for B) was injected after 60 seconds and washed with HBS after 200 seconds for an additional 600 seconds (for A) and 200 seconds (for B). Biacore T100 Evaluation software version 1.1.1 was used.

5.3.5.4 Binding hPC107 to wild type and mutant HCS

In this experiment, immobilised wild type and mutant HCS were pre-charged with 125 µM biotin:MgATP and the kinetics of apo hPC107 binding determined. This was performed in order to determine the effect of the MCD mutants on the association and dissociation rates of hPC107 binding to HCS. Figure 5.10 shows a representative SPR chromatogram from three independent experiments when 20 µM hPC107 was included in the binding buffer. There was no response in the blank flow cell when hPC107 was injected and therefore hPC107 did not bind to the chip non-specifically.

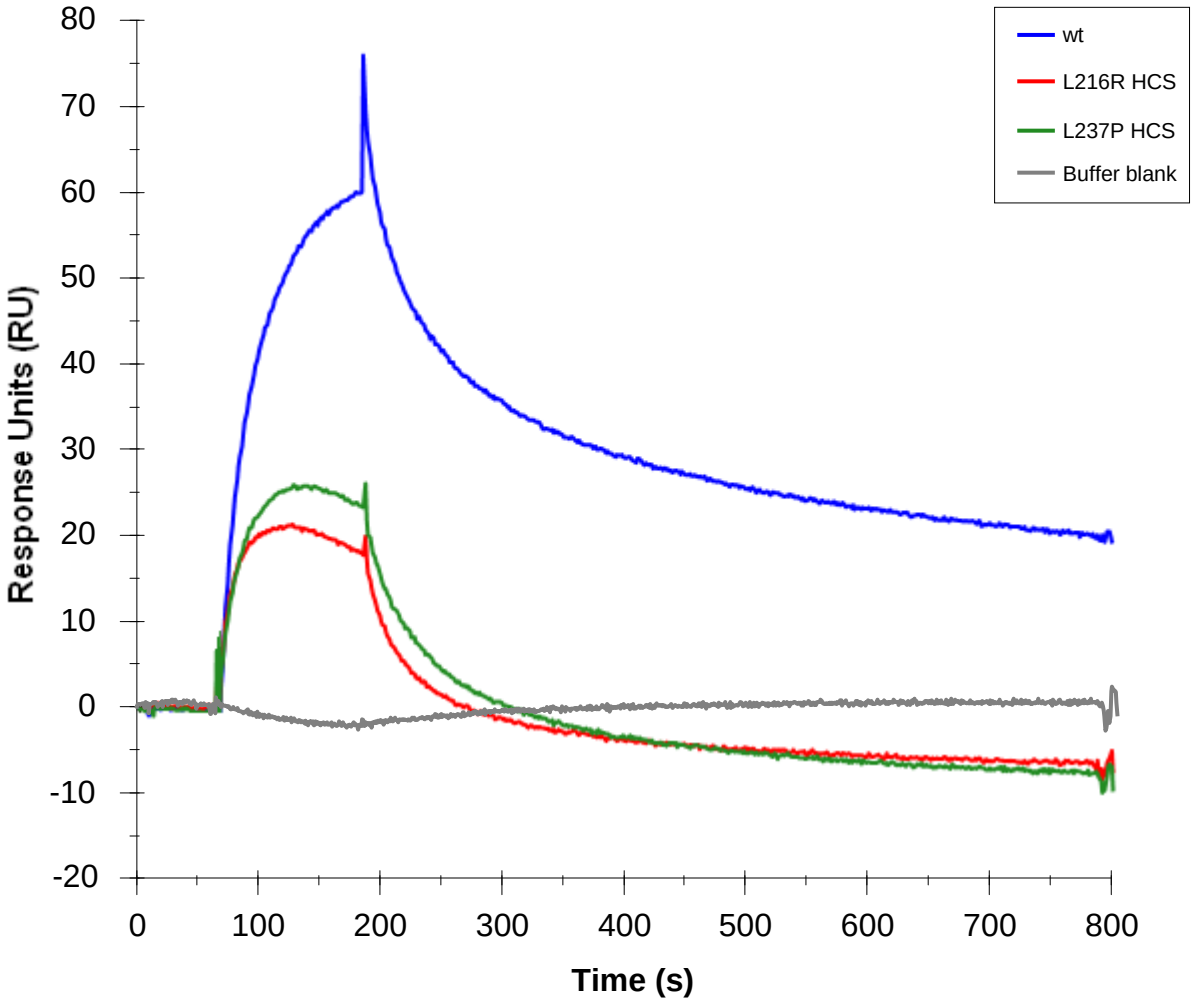


Figure 5.10 Sensorgram showing the association and dissociation of apo hPC107 to and from wild type and MCD mutant HCS. Enzymes pre-charged with biotin:MgATP were injected with 20 µM hPC107 and after 200 seconds were washed with HBS for a further 600 seconds. Biacore T100 Evaluation software version 1.1.1 was used.

The higher amplitude in response units observed with wild type HCS were due to the fact that at 20 µM hPC107 saturating conditions have been reached as the equilibrium constant, K_D was 1.0 ± 0.7µM (table 5.1). This was consistent with the measured K_M of 1.0 ± 0.2 µM determined using the *in vitro* biotinylation assay (Polyak, unpublished data). The mutant HCS K_D values for hPC107 however were between 16-19 µM (table 5.1) and therefore saturating conditions were not reached. Residual hPC107 remained bound to wild type HCS for 800 seconds as the response did not return to zero. Although the

response was declining it will require more time for complete dissociation to occur. For the mutant HCS, the curve decreases below zero, which is indicative of mutant HCS discharging the biotinyl-5´-AMP. As figure 5.10 shows, the maximum response for biotin:ATP was 25-35 RU for the HCS mutants. For complete loss of substrate to be observed, the curve would be expected to decrease below zero by 25-35 RU. For the HCS mutants hPC107 was completely unbound after 350 seconds as they presumably have completed their catalytic cycle.

Table 5.1 summarises the determined association rate constant (k_a), dissociation rate constant (k_d) and the equilibrium dissociation constant (K_D) for the data presented in figure 5.9. The association rates of hPC107 with wild type and MCD mutant HCS were comparable. Wild type HCS had a K_D of 1.0 ± 0.7µM. However, for the mutant HCS, the dissociation rates were 16-18 times faster than wild type. This resulted in elevations of the K_D for the mutant HCS for its substrate. The K_D values for L216R and L237P HCS were 18.9 and 16.9 µM, respectively, significantly higher that wild type ($p<0.01$).

Table 5.1 Kinetic and equilibrium constants determined for the interaction of HCS with hPC107 at pH 7.4 and 25°C

Enzyme	**k_a (M^{-1} s^{-1}) x10^{-3}**	**k_d (s^{-1}) x 10^3**	**K_D (µM)**
Wild type HCS	1851 ± 17	0.002 ± 0.0003	1.03 ± 0.74
L216R HCS	1843 ± 23	0.035 ± 0.0055	18.9 ± 4.80
L237P HCS	1852 ± 6	0.031 ± 0.0085	16.9 ± 4.14

Note: The values presented are means ± standard deviations of three independent assays.

Using the Biacore T100 Evaluation software version 1.1.1 data for hPC-107 binding was fitted to a 1:1 ligand-binding model. The goodness of fit is indicated by the Chi^2 value, which were 0.25, 0.43 and 0.51 for wild type HCS, L216R HCS and L237P HCS, respectively. Since the Chi^2 values are lower than 10% of the calculated maximum response (they are <2%) the curves fit the binding model well. Residuals are the difference between the experimental and calculated data points and provide another measure to

determine the relative goodness of fit for the model. If the kinetic model sufficiently describes the interactions between HCS and hPC107, the difference between the calculated and experimental points will be minimal and the residuals will be randomly distributed around a zero value. The residual plots (figure 5.11) for the fitting show minimal discrepancies between the experimental data and the fit.

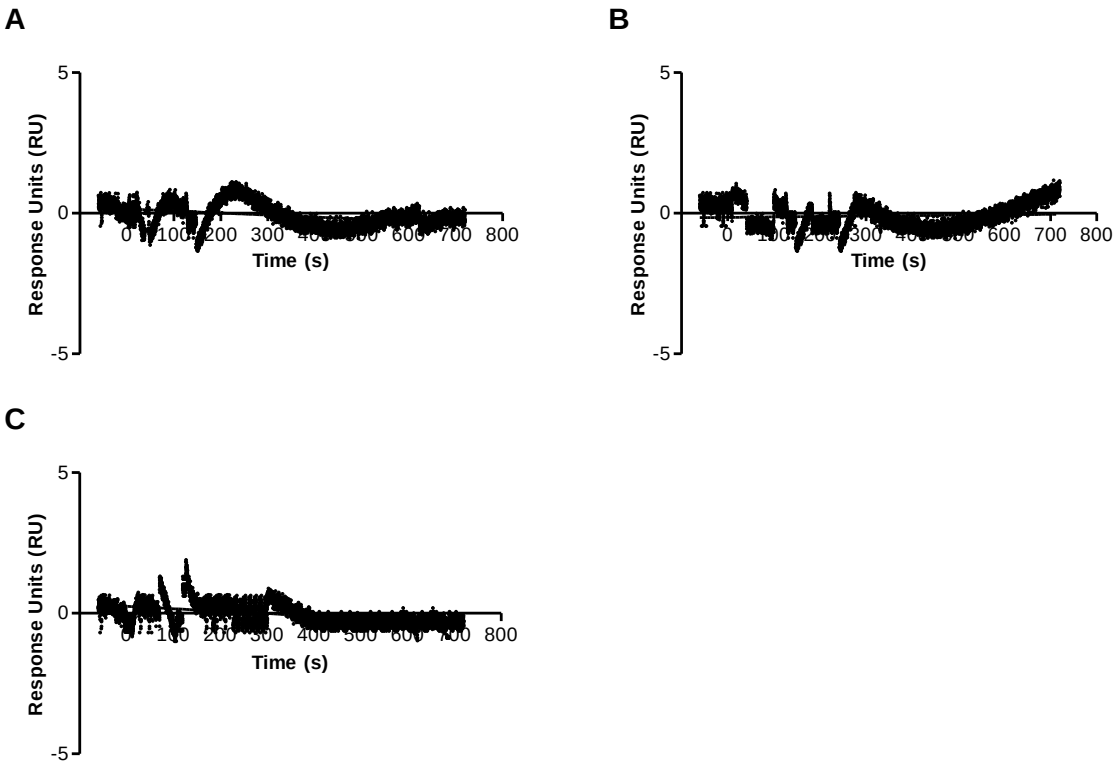


Figure 5.11 Residual plots of the binding of hPC107 to **(A)** wild type HCS, **(B)** L216R HCS and **(C)** L237P HCS using the 1:1 binding model.

It was also determined whether pre-charging with 125 µM biotin:MgATP was required for the binding of hPC107 to HCS (figure 5.12). It was found using SPR that pre-charging was not necessary to determine the binding of hPC107 to HCS.

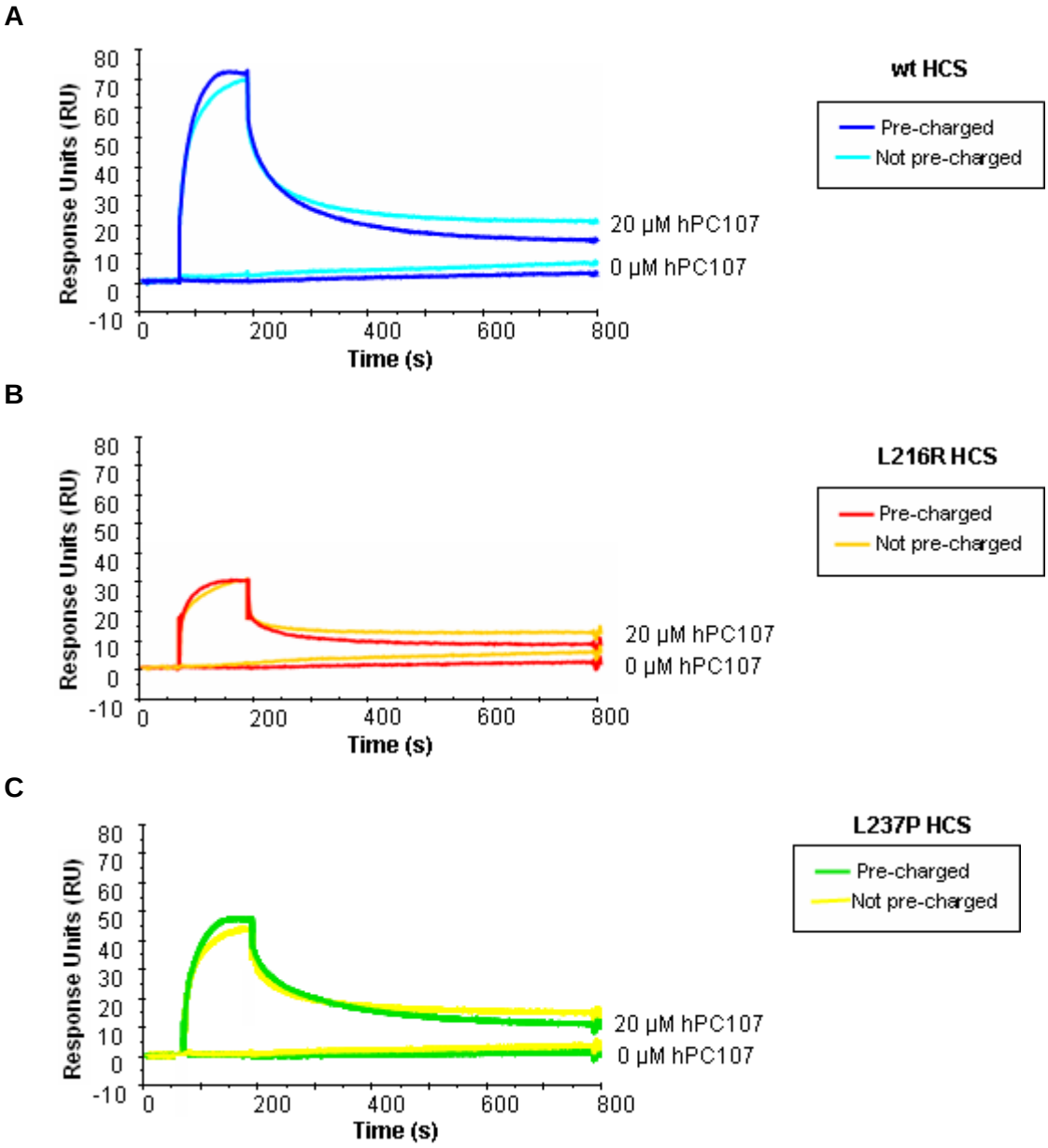


Figure 5.12 Sensorgrams showing the association and dissociation of apo hPC107 to and from **(A)** wild type, **(B)** L216R HCS and **(C)** L237P HCS. Enzymes were either pre-charged or not pre-charged with biotin:MgATP and were injected with 20 µM hPC107 and after 200 seconds were washed with HBS for a further 600 seconds. Biacore T100 Evaluation software version 1.1.1 was used.

5.4 Discussion

MCD is treated routinely by the administration of biotin. There is however a group of HCS-deficient patients that respond poorly to oral biotin therapy and the HCS mutants that give rise to MCD in this group of patients contain mutations in the N-terminal domain of HCS. These mutations, namely R183P (no clinical data available), L216R and L237P HCS, were found to have a significant decrease in V_{MAX}, however, a similar K_M for biotin was reported (Aoki *et al.*, 1997; Sakamoto *et al.*, 1998). Yokoi *et al.* (2009) found that the L237P HCS mutant could not be treated with prenatal biotin therapy. This Chapter aimed to further characterise the mechanism for biotin non-responsiveness of the V_{MAX} MCD mutants and in doing so functionally characterise the N-terminal region of HCS.

The binding affinities of apo hPC107 for wild type and MCD mutant HCS was determined using Biacore™. The association rates of hPC107 with wild type and MCD mutant HCS were comparable as they contained similar K_a values. The dissociation rates of hPC107 from L216R and L237P HCS were 16-18 times faster than wild type HCS and I conclude the MCD mutant HCS does not stay bound to its substrate long enough for biotin transfer to occur. The data presented in this Chapter shows that the the observation of a lack of interaction between MCD mutant HCS and hPC107 in Chapter 3 was most likely due to the dissociation being too rapid for the GAL1 transcription reporter gene to be switched on.

In Chapter 3 it was hypothesised that the second partial reaction of HCS is affected by the V_{MAX} MCD mutations as the interaction between HCS and the biotin domain occurs in the second partial reaction. The Yeast Two-hybrid assay showed that the MCD mutants disrupted the interaction between HCS and hPC107. This presented a novel mechanism where the second partial reaction is affected by the MCD HCS mutants at the qualitative level. In this Chapter SPR evidence supported the proposal that the MCD mutants are not

compromised in biotinyl-5´-AMP synthetase activity. It was shown that injecting immobilised HCS with biotin and MgATP forms a slow dissociating species, which was presumed to be biotinyl-5´-AMP as this species has the characteristics of biotinyl-5´-AMP described by Xu and Beckett (1994). This also agrees with Aoki *et al.* (1997) who found that the L237P HCS mutant had a similar K_M for biotin as the wild type indicating the first partial reaction is potentially not affected by the mutation, which is consistent with our data.

Further quantitative characterisation of the disruption of the interaction between HCS and hPC107 was performed using SPR (Biacore^{TM}). Once pure protein was obtained, *in vitro* HCS assays were performed on each enzyme. Using hPC107 as substrate, the specific activity of L216R HCS and L237P HCS were 3.3 % and 3.5 %, of wild type, respectively. This was in excellent agreement with Aoki *et al.* (1997) who determined the HCS activities of patients harbouring the L237P HCS mutant decreased to 4.0 % and 1.6 % for the two patient lymphoblasts assayed. It was shown using Circular Dichroism that the MCD mutants do not effect the gross overall fold of HCS as proteins contained overlapping spectra. However, the stability of L216R HCS has been shown to be affected (Bailey *et al.*, 2008). In a cell culture model the L216R HCS mutant displayed a turn over rate double that of wild type HCS. Although gross overall fold of HCS was not affected by the mutations, the stability of the enzyme is compromised through a mechanism that is not known.

The characterisation of the interactions of wild type and mutant HCS were performed by immobilising the purified wild type, L216R and L237P HCS to the surface of a CM5 sensor chip. The first experiment that was conducted was to verify biological activity of the immobilised enzymes. This was performed by injecting increasing concentrations of biotin onto the enzyme bound chip. It was shown that MgATP also binds to the enzymes,

however the response units (RU) were lower than that for biotin binding. This correlates with the Ingaramo and Beckett (2009) study, which indicated that biotin binds first, followed by ATP. In this Chapter it was shown that when biotin and MgATP were injected together, the synergistic effect resulted in a 6 fold increase in RU compared with biotin alone. The resulting sensorgram of biotin:ATP binding showed a slow dissociating species was formed in wild type and MCD mutant HCS indicating that biotinyl-5´-AMP was potentially synthesised as the species behaved as described by Xu and Beckett (1994). This data along with the fact that biotin binds to the MCD mutant HCS proteins supports the hypothesis that the first partial reaction is occurring in these mutants. Interestingly, this study has shown that biotinol-5´-AMP does not form a slow dissociating species in HCS indicating that the selective inhibition for *S. aureus* (Soares da Costa, personal communication) that has been observed using this compound is due to a rapid dissociation from the active site of HCS. This data indicates that the N-terminal region of HCS is a feature that may be beneficial in the discovery of novel therapeutics as it may have a role in substrate specificity. In Chapter 3 it was shown that the N-terminal cap (NTC) was able to interact with C-HCS and may function as cap closing over the active site, possibly to help stabilise the reaction intermediate from diffusing out of the active site. This Chapter showed that this cap may also function in locking the biotin domain of biotin-dependent enzymes to the C-terminal region of HCS during biotin transfer.

The mechanism for disease that I presented here is that while the MCD mutant enzymes are associating with the substrate at a comparable rate to wild type HCS it is the increased dissociation rate that leads to insufficient biotinylation of the substrate. That is, the dissociation rates are too rapid for efficient biotinylation to occur. It was also shown that the first partial reaction is likely occurring and this is why biotin supplementation is not a suitable treatment for disease with the V_{MAX} MCD mutant cases.

CHAPTER 6

Final Discussion

6.1 Functional characterisation of the N-terminal region of HCS

The work presented in this thesis showed that the N-terminal domain interacts with the catalytic/C-terminal domain. This thesis also provided a function for the proposed structural N-terminal domain of HCS, NTC. NTC is indirectly involved the interaction between HCS and its substrate hPC107 as this interaction was disrupted by the NTC mutations. A recent study by Lee *et al.* (2009) showed that residues M1-L160 of HCS interacted with hACC75 (human Acetyl CoA Carboxylase 75) using NMR titration. A study by Hassan *et al.* (2009) found that residues 1-446 N-terminal HCS interacted with p67 (propionyl-CoA carboxylase 67), the interaction was qualitatively described as weak, given the long incubation times required for the yeast to grow. Another recent study highlighting the importance of the N-terminal region of HCS in substrate recognition was performed by Ingaramo and Beckett (2009). Their study kinetically characterised the two isoforms M1 and M58 HCS. They found that M1 had a slightly higher k_{cat} than M58. Importantly, the study found that Met1 HCS associates with the model substrate p67 (propionyl-CoA carboxylase 67) two fold faster than Met58 HCS, therefore implicating the N-terminal region of HCS in substrate recognition. Although the focus of the work presented in this thesis looks at mutations that lie between G159-M315 and not M1-L160, as in the previous mentioned studies, it also shows the functional importance of N-terminal HCS in substrate recognition.

Together with studies from Lee *et al.* (2009), Hassan *et al.* (2009) and Ingaramo and Beckett (2009), the work in this thesis indicates that N-terminal HCS is required for substrate binding and is involved in stabilising the enzyme:substrate complex. This study does however look at specific residues in the N-terminal region of HCS that may be interacting with the protein substrate and are important in the formation of this interaction.

6.2 A novel molecular mechanism to explain the juvenile metabolic syndrome biotin-unresponsive HCS deficiency

This thesis also gives an explanation for the lack of biotin responsiveness of the MCD V_{MAX} patients. Patients that contain the L216R and L237P HCS mutations are not responsive to biotin therapy because the first partial reaction of HCS is occurring and therefore it is not the limiting factor. It was observed in this study that the reaction intermediate biotinyl-5´-AMP is formed by the catalytically active MCD mutants. This work supported the second hypothesis as it was shown that the second partial reaction of HCS is affected by the MCD mutations. Initially, the Yeast Two-hybrid assay was used to determine the effect of the mutations on the interaction of HCS and its substrate. The interaction was disrupted by the MCD mutations and SPR was used to characterise this observation further. I showed that although the association between HCS and substrate was occurring at a similar rate, the MCD mutants were dissociating from the substrate at a much faster rate than wild type enzyme. This shows that the MCD mutations prevent efficient biotinylation of the biotin-dependent enzymes due to the rapid dissociation rates of mutant HCS from the substrates. The work in this thesis explained why the Yokoi *et al.* (2009) study found that the L237P HCS mutant could not be treated with prenatal biotin therapy. Whether biotin is administered prenatally or after birth, the V_{MAX} mutations cannot be treated in this manner as it is the interaction of HCS with its protein substrate that is affected. This highlights that alternative treatments are required. Due to this, alternative therapies to biotin supplementation are required for the treatment of patients that are not responsive to the therapy.

In this study I was able to determine the K_D of hPC107 for wild type and MCD mutant HCS using SPR analysis. In future studies this technique can be used to test the hypothesis that HCS selects between enzyme substrates. A hierarchy of biotin-dependent enzymes can

be drawn from the K_D values obtained. It would also be interesting to probe this further by delineating the effect of the MCD mutant HCS on the K_D of each biotin-dependent enzyme.

6.3 New directions for the development of V_{MAX} HCS deficiency therapeutics

HCS allosteric activators

A potential therapeutic for MCD is the administration of HCS allosteric activators. Allosteric activators increase enzyme activity by binding at an allosteric site that affects turnover at the catalytic site. An example is the allosteric activation of carbonic anhydrase. Carbonic anhydrase catalyses the reversible reaction between CO_2 hydration and HCO_3^- dehydration and activation of this enzyme provides a rapid mechanism to raise HCO_3^- concentrations in synapses. The enzyme functions as a gate that controls signal transfer through the neural network. Allosteric activators have been successful in the activation of the zinc enzyme carbonic anhydrase isoforms hCA I, II and IV with L-histidine and some of its derivatives (Sun and Alkon, 2002; Abdo *et al.*, 2009). This approach has been shown to enhance the synaptic efficacy, which may facilitate a treatment of Alzheimer's disease, aging and other conditions involving learning and memory therapy (Sun and Alkon, 2001; Sun and Alkon; 2002). L-histidine derivative carbonic anhydrase activators continue to investigated, recently (Abdo *et al.*, 2009) an isoform selective potent carbonic anhydrase activator was found.

For a structure based approach to work it is important to have a crystal structure of HCS, as any possible allosteric activators using the enzyme assay would be required to be docked or co-crystallised with HCS in order to improve upon their activation potential. An alternative approach to structure based drug design is performing a chemical library screen. Work with carbonic hydrate activators has been drastically enhanced by the solving of the X-ray crystal structure with L-histidine as this work showed that the moieties

of L-Histidine do not participate in interactions with amino acid moieties of the active site and can be derivatised to obtain potent activators (Temperini *et al.*, 2005). Currently a high resolution structure of HCS remains to be elucidated; solving the structure would aid in the development of HCS allosteric activators.

Another example where allosteric activators were explored as a therapy is diabetes. Glucokinase maintains plasma glucose homeostasis by enhancing insulin secrection from pancreatic β-cells and glucose metabolism in the liver (Nishimura *et al.*, 2009). Mutations in glucokinase lead to maturity onset diabetes of the young type 2 (MODY2). The mutations result in a decreased V_{MAX} of the enzyme and reduced affinity for its substrates glucose and ATP (Grimsby *et al.*, 2003). Grimsby *et al.* (2003) searched for allosteric activators of glucokinase by screening a structurally diverse library of 120 000 compounds. This screening process led to one hit compound that increased the enzymatic activity of glucokinase, which was chemically optimised. They found that this glucokinase activator increased the affinity for glucose and the V_{MAX}. In rodents, the glucokinase activator lowered blood glucose levels and increased hepatic glucose uptake. This study showed that intracellular enzymes can be targeted and that the development of allosteric activators can potentially lead to new drug therapies. There have been recent developments such as the Nishimura *et al.* (2009) study, which developed potent novel 2-amino benzamide glucokinase activators. One of the compounds demonstrated *in vivo* glucose-lowering effects in an acute rat model.

Evidence of activators in biotin protein ligase has only been shown with the *E. coli* BPL (BirA), where biotin acts as an activator of the biotin repressor function of BirA (Weaver *et al.*, 2001). No evidence of allosteric activation of HCS by any compound has been shown. However, screening for activators from a large library as performed in the Grimsby *et al.*

(2003) study is a starting point. This approach may be time consuming as the widely used ^{3}H-biotin incorporation assay is not well suited to high throughput assays. However, to overcome this, our lab has been developing a plate-based assay that could potentially be employed in high throughput applications.

Enzyme replacement therapy

A potential therapy for HCS deficiency is enzyme replacement therapy (ERT), which is an approach that replaces the inactive enzyme in affected patients via intravenous infusion with functional recombinant enzyme or enzyme purified from tissue such as placenta. ERT has been successful for the treatment of Gaucher Type 1 (caused by a genetic deficiency of lysosomal glucocerebrosidase) by the administration of alglucerase or imiglucerase (Lin *et al.*, 2006). ERT has been successful in treating Anderson-Fabry disease, which is an X-linked disorder that is caused by a deficiency in lysosomal α-galactosidase A, which results in the inability to clear glycosphingolipids. Recombinant human α-galactosidase A was administered and successfully treated the disease (Basic-Jukic *et al.*, 2009). Another example of ERT success is the treatment of Pompe disease. Pompe disease is due to mutations in α-glucosidase, which is a glycogen-degrading lysosomal enzyme. Recombinant human α-glucosidase was used to successfully treat the disease (Sun *et al.*, 2007). These examples are all intra-cellular enzymes as is HCS and the success of enzyme replacement therapy in lysosomal enzymes is encouraging as a potential therapy for MCD. However, lysosomes are able to readily import foreign material by endocytosis and therefore more suited to ERT than the cytosol or mitochondra where HCS is located. Delivery of the enzymes for ERT in the case of HCS therefore may be challenging. Examples of delivery strategies are cloning a mitochondrial targeting signal at the N-terminus of HCS for delivery of HCS into mitochondria and using peptide-mediated cellular transduction.

Small molecular chaperones

An emerging approach to treating protein deficiencies is pharmacological chaperone therapy. This differs from the allosteric activators as these are active-site-specific chaperones. These chaperones target the enzyme active-site directed by small molecule pharmacological chaperones serving as a folding template that can dramatically shift the folding dynamics in favor of proper and native-like folding. This prevents the degradation of mutant proteins in the endoplasmic reticulum (ER) (Fan, 2008). An example where small molecule pharmacological chaperones were used is the Flanagan *et al.* (2009) study were they found the pharmacological chaperone 1-deoxynojirimycin (DNJ) increased the activity and lysosomal trafficking of multiple mutant forms of acid α-glucosidase, which cause Pompe disease. The study characterised the pharmacological chaperone DNJ on 76 different mutant forms of α-glucosidase identified in Pompe disease. The chaperone significantly increased enzyme activity and protein levels for 16 different α-glucosidase mutants in patient-derived fibroblasts. DNJ also increased the processing of the α-glucosidase mutants to their mature lysosomal forms, which suggested that the chaperone assisted trafficking through the secretory pathway. They conducted immunofluorescence microscopy studies that showed increased colocalisation of α-glucosidase with LAMP2, the lysosomal marker, after incubation with DNJ. This also confirmed increased lysosomal trafficking. The data supported further evaluation of DNJ as a potential treatment for Pompe disease.

Bailey *et al.* (2008) showed that the stability of L216R HCS was affected as the turn over rate was double that of wild type HCS. Although gross changes in secondary structure were not affected (CD data from this study), the stability of the enzyme was compromised. Small molecular chaperones could be used as a treatment to prevent the instability of the V_{MAX} MCD mutants as the chaperones would serve as a folding template and result in

native-like folding. This would prevent the degradation of V_{MAX} HCS and thereby to some extent treat the biotin unresponsive MCD patients.

I have presented several potential therapies for MCD patients that are not responsive to biotin therapy. Some of the suggested therapies are in the infant stages of development and have not been attempted for HCS. However, small molecular chaperones, HCS allosteric activators and enzyme replacement therapy may become alternatives to biotin therapy. Realistically, having a viable treatment this is only a possibility in the long term.

References

Abdo, M., Vullo, D., Saada, M., Montero, J., Scozzafava, A., Winum, J. and Supuran, C.T. (2009) Carbonic anhydrase activators: Activation of human isozymes I, II an dIX with phenylsulfonylhydrazido L-histidine derivatives, *Bioorganic & Med Chem Let*, 19, 2440-2443.

Ames, B. N., Elson-Schwab, I. and Silver, E. A. (2002) High-dose vitamin therapy stimulates variant enzymes with decreased coenzyme binding affinity (increased K_M): Relevance to genetic disease and polymorphisms, *Am J Clin Nutr,* 75**,** 616-58.

Aoki, Y., Li, X., Sakamoto, O., Hiratsuka, M., Akaishi, H., Xu, L., Briones, P., Suormala, T., Baumgartner, E. R., Suzuki, Y. and Narisawa, K. (1999) Identification and characterization of mutations in patients with holocarboxylase synthetase deficiency, *Hum Genet,* 104**,** 143-8.

Aoki, Y., Suzuki, Y., Li, X., Sakamoto, O., Chikaoka, H., Takita, S. and Narisawa, K. (1997) Characterization of mutant holocarboxylase synthetase (HCS): A *KM* for biotin was not elevated in a patient with HCS deficiency, *Pediatr Res,* 42**,** 849-54.

Aoki, Y., Suzuki, Y., Sakamoto, O., Li, X., Takahashi, K., Ohtake, A., Sakuta, R., Ohura, T., Miyabayashi, S. and Narisawa, K. (1995) Molecular analysis of holocarboxylase synthetase deficiency: A missense mutation and a single base deletion are predominant in Japanese patients, *Biochim Biophys Acta,* 1272**,** 168-74.

Athappilly, F. K. and Hendrickson, W. A. (1995) Structure of the biotinyl domain of acetylcoenzyme A carboxylase determined by MAD phasing, *Structure*, 3, 1407-19.

Bailey, L.M., Wallace, J.C. and Polyak, S.W. (2010) Holocarboxylase synthetase: Correlation of protein localisation with biological function, *Archives Biochem & Biophys*, 496, 45-52.

Bailey, L.M., Ivanov, R.A., Jitrapakdee, S., Wilson, C.J., Wallace, J.C. and Polyak, S.W. (2008) Reduced Half-life of Holocarboxylase Synthetase from Patients with Severe Multiple Carboxylase Deficiency, *Hum Mut*: *Mut in Brief*, 29, E47-E57.

Bailey, L.M., Ivanov, R.A., Wallace, J.C. and Polyak, S.W. (2008) Artifactual detection of biotin on histones by streptavidin, *Anal Biochem*, 373, 71-77.

Bagautdinov, B., Kuroishi, C., Sugahara, M. and Kunishima, N. (2005) Crystal structures of biotin protein ligase from *Pyrococcus horikoshii* OT3 and its complexes: Structural basis of biotin activation, *J Mol Biol*, 353, 322-33.

Bagautdinov, B., Matsuura, Y., Bagautdinova, S. and Kunishima, N. (2008) Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate, *J Biol Chem*, 283, 14739-14750.

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

Bartel, P., Chien, C. T., Sternglanz, R. and Fields, S. (1993) Elimination of false positives that arise in using the two-hybrid system, *Biotechniques*, 14, 920-4.

Basic-Jukic, N., Kes, P., Mokos, I. and Coric, M. (2009) Do we need more intensive enzyme replacement therapy for Anderson-Fabry disease, *Med Hypotheses*, 473-483.

Baumgartner, E. R. and Suormala, T. (1997) Multiple carboxylase deficiency: Inherited and acquired disorders of biotin metabolism, *Int J Vitam Nutr Res*, 67, 377-84.

Black, S., Wilcock, G.K., Hawworth, J., Hendrix, S., Zavitz, K., Christensen, D.B., Bass, S., Laughlin, M., and Swabb, E. (2006) Efficacy and safety of MPC-7869 (R-flurbiprofen), a selective Abeta42 lowering agent in mild Alzheimer's disease: results of a 12-month phase 2 trial and 1-year follow on study, *Neurology*, 66, Suppl2.

Burri, B. J., Sweetman, L. and Nyhan, W. L. (1981) Mutant holocarboxylase synthetase: Evidence for the enzyme defect in early infantile biotin-responsive multiple carboxylase deficiency, *J Clin Invest,* 68**,** 1491-5.

Burri, B. J., Sweetman, L. and Nyhan, W. L. (1985) Heterogeneity of Holocarboxylase synthetase in patients with biotin-responsive multiple carboxylase deficiency, *Am J Hum Genet,* 37**,** 326-37.

Brown, P.H. and Beckett, D. (2005) Use of Binding Enthalpy to Drive an Allosteric Transition, *Biochemistry*, 44, 3112-3121.

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

Cai, J., Huang, Y., Li, F. and Li, Y. (2006) Alteration of protein subcellular location and domain formation by alternative translational initiation, *Proteins,* 62**,** 793-9.

Campeau, E. and Gravel, R. A. (2001) Expression in Escherichia coli of N- and C-terminally deleted human holocarboxylase synthetase. Influence of the N-terminus on biotinylation and identification of a minimum functional protein, *J Biol Chem,* 276**,** 12310-6.

Chang, H.I. and 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. and Cronan, J. E., Jr. (1999) The enzymatic biotinylation of proteins: A post-translational modification of exceptional specificity, *Trends Biochem Sci*,24, 359-63.

Chapman-Smith, A. and Cronan, J. E., Jr. (1999) Molecular biology of biotin attachment to proteins, *J Nutr,* 129**,** 477-484.

Chapman-Smith, A., Morris, T. W., Wallace, J. C. and Cronan, J. E., Jr. (1999) Molecular recognition in a post-translational modification of exceptional specificity. Mutants of the biotinylated domain of acetyl-CoA carboxylase defective in recognition by biotin protein ligase, *J Biol Chem,* 274**,** 1449-57.

Chen, H. (2008) Ca^{2+} and Phosphoinositides Regulations in α-actinin-4 F-actin Binding, PhD Thesis, University of Adelaide, Adelaide.

Choi-Rhee, E., Schulman, H. and Cronan, J.E. (2004) Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase, *Protein Science*, 13, 3043-3050.

Christensen, Q.H. and Cronan, J.E. (2010) Lipoic acid synthesis: a new family of octanoyltransferases generally annotated as lipoate protein ligases, *Biochemistry*, 49, 10024-10036.

CLONTECH Laboratories. (1996) User Manual (PT3040-1) MATCHMAKER LexA Two-Hybrid System, Palo Alto, CA.

Clowes, R.T., Crawford, A., Raine, A.R., Smith, B.O. and Laue, E.D. (1995) Improved methods for structural studies of proteins using nuclear magnetic resonance spectroscopy, *Cur Opin Biotechnol*, 6, 81-88.

Cohen, N. D., Thomas, M. and Stack, M. (1985) The subcellular distribution of the holocarboxylase synthetase of the rat liver, *Annals of the New York Academy of Sciences,* 447**,** 393-395.

Cronan, J. E., Jr. (1995) The E. coli bio operon: Transcriptional repression by an essential protein modification enzyme, *Cell,* 58**,** 427-9.

Dotsch, V., and Wagner, G. (1998) New approaches to structure determination by NMR Spectroscopy, *Cur opinion Struct Biol*, *8*, 619-623.

Dupuis, L., Campeau, E., Leclerc, D. and Gravel, R. A. (1999) Mechanism of biotin responsiveness in biotin-responsive multiple carboxylase deficiency, *Mol Genet Metab,* 66**,** 80-90.

Dupuis, L., Leon-Del-Rio, A., Leclerc, D., Campeau, E., Sweetman, L., Saudubray, J. M., Herman, G., Gibson, K. M. and Gravel, R. A. (1996) Clustering of mutations in the biotin-binding region of holocarboxylase synthetase in biotin-responsive multiple carboxylase deficiency, *Hum Mol Genet,* 5**,** 1011-6.

Eswar, N., Eramian, D., Webb, B., Shen, M. and Sali, A. (2006) Protein structure modeling with MODELLER, *Methods in Mol Biol*, 426, 145-159.

Fan, J. Q. (2008) A counterintuitive approach to treat enzyme deficiencies: Use of enzyme inhibitors for restoring mutant enzyme activity, *Biol Chem,* 389**,** 1-11.

Feldman, G. L., Hsia, Y. E. and Wolf, B. (1981) Biochemical characterization of biotinresponsive multiple carboxylase deficiency: Heterogeneity within the bio genetic complementation group, *Am J Hum Genet,* 33**,** 692-701.

Fernandez, C. and Wider, G. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules, *Current Opinion Struct Biol*, 13, 570-580.

Flanagan, J.J., Rossi, B., Tang, K., Wu, X., Mascioli, K., Donaudy, F., Tuzzi, M.R., Fontana, F., Cubellis, M.V., Porto, C., Benjamin, E., Lockhart, D.J., Valenzano, K.J., Andria, G., Parenti, G., Do, H.V. (2009) The pharmacological chaperone 1-deoxynojirimycin increases the activity and lysosomal trafficking of multiple mutant forms of acid alpha-glucosidase, *Hum Mut*, 30, 1683-1692.

Gardner, K.H. and Kay, L.E. (1998) The use of ^{2}H,^{13}C,^{15}N multidimensional NMR to study the structure and dynamics of proteins, *Annu Rev Biophys Biomol Struct,* 27, 357-406.

Georgakopoulou, S., Möller, D., Sachs, N., Herrmann,H. and Aebi, U. (2009) Near-UV Circular Dichroism Reveals Structural Transitions of Vimentin Subunits during Intermediate Filament Assembly, *J Mol Biol*, 386, 544-553.

Greenfield, N. and Fasman, G.D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation, *Biochemistry*, 8, 4108–4116.

Grimsby, J., Sarabu, R., Corbett, W.L., Haynes, N-E, Bizzarro, F.T.,Coffey, J.W.,Guertin, K.R., Hilliard, D.W., Kester, R.F., Mahaney, P.E., Marcus, L., Qi, L., Spence, C.L., Tengi, J., Magnuson, M.A., Chu, C.A., Dvorozniak, M.T., Matschinsky, F.M. and Grippo, J.F. (2003) Allosteric Activators of Glucokinase: Potential Role in Diabetes Therapy, *Science*, 301, 370-373.

Hassan, Y.I., Moriyama, H., Olsen, L.J., Bi, X. and Zempleni, J. (2009) N- and C-terminal domains in human Holocarboxylase synthetase participate in substrate recognition, *Mol Gen & Metabol*, 96, 183-188.

Healy, S., McDonald, M.K., Wu, X., Yue, W.W., Kochan, G., Oppermann, U. and Gravel R.A. (2010) Structural impact of human and Escherichia coli biotin carboxyl carrier proteins on biotin attachment, *Biochemistry*, 49, 4687-4694.

Healy, S., Perez-Cahahia, B, Jia, D., McDonald, M.G., Davie, J.R. and Gravel, R,A. (2009) Biotin is not a natural histone modification, *Biochimica et Biophysica Acta*, 719-733.

Healy, S., Heightman, T.D., Hohmann, L., Schriemer, D. and Gravel, R.A. (2009) Nonenzymatic biotinylation of histone H2A, *Prot Sci*, 18, 314-328.

Hennessey, J. P., and Johnson, W. C. (1981) Information content in the circular dichroism of proteins. *Biochemistry*, 20, 1085–1094.

Hiratsuka, M., Sakamoto, O., Li, X., Suzuki, Y., Aoki, Y. and Narisawa, K. (1998) Identification of holocarboxylase synthetase (HCS) proteins in human placenta, *Biochim Biophys Acta,* 1385**,** 165-71.

Holzwarth, G. and doty, P. (1965) The ultraviolet circular dichroism of polypeptides, *J Am Chem Soc*,87, 218-228.

Howard, M. J. (1998) Protein NMR spectroscopy, *Cur Biology*, 8, R331-R333.

Hymes, J., Fleischhauer, K. and Wolf, B. (1995) Biotinylation of histones by human serum biotinidase: Assessment of biotinyl-transferase activity in sera from normal individuals and children with biotinidase deficiency, *Biochem Mol Med,* 56**,** 76-83.

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

Ingaramo, M. and Beckett, D. (2009) The distinct N-termini of two human HCS isoforms influence biotin acceptor substrate recognition, *JBC*, 1-21.

Johnson, W.C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures, *Proteins: Struct Funct Genet,* 35, 307–312.

Kelley, L.A. and Sternberg, M.J.E. (2009) Protein structure prediction on the Web: a case study using the Phyre server, *Nat Prot*, 4, 363-371.

Kleywegt, G.J. and Jones, T.A. (1996) Phi/Psi-chology: Ramachandran revisited, *Structure*, 4, 1395-1400.

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

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

Kukar, T. and Golde, T.E. (2008) Possible Mechanisms of Action of NSAIDs and Related Compounds that Modulate γ-Secretase Cleavage, *Cur Top Med Chem*, 8, 47-53.

Lee, C., Cheong, C. and Jeon, Y.H. (2010) The N-terminal domain of human holocarboxylase synthetase facilitates biotinylation via direct interaction with the substrate protein, *FEBS Letters*, 584, 675-680.

Lee, C., Cheong, C. and Jeon, Y.H. (2009) Substrate recognition characterists of human Holocarboxylase synthetase for biotin ligation, *Biochem & Biophys Res Commun*, 391, 455-460.

Leon-Del-Rio, A., Leclerc, D., Akerman, B., Wakamatsu, N. and 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,* 92**,** 4626-30.

Liedberg, B., Nylander, C. and Lunström, I. (1983) Surface plasmond resonance for gas detection and biosensing, *Sensors & Actuators*, 4, 299-304.

Lin, H., Lin, S., Chuang, C. And Wraith, J.E. (2006) Enzyme replacement therapy with imiglucerase in a Taiwanese child with type 1 Gaucher disease, *J Chin Med Assoc*, 69, 228-232.

McKean, A.L., Ke, J., Song, J., Che, P., Achenbach, S., Nikolau, B.J. and Wurtele, E.S. (2000) Molecular characterization of the non-biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase, *J Biol Chem*, 275,5582-5590.

Morrone, A., Malvagia, S., Donati, M. A., Funghini, S., Ciani, F., Pela, I., Boneh, A., Peters, H., Pasquini, E. and Zammarchi, E. (2002) Clinical findings and biochemical and molecular analysis of four patients with Holocarboxylase synthetase deficiency, *Am J Med Genet,* 111**,** 10-8.

Naganathan, S. and Beckett, D. (2007) Nucleation of an allosteric response via ligand-induced loop folding, *J Mol Biol,* 373, 96-111.

Narang, M. A., Dumas, R., Ayer, L. M. and Gravel, R. A. (2004) Reduced histone biotinylation in multiple carboxylase deficiency patients: A nuclear role for holocarboxylase synthetase, *Hum Mol Genet,* 13**,** 15-23.

Narisawa, K., Arai, N., Igarashi, Y., Satoh, T., Tada and K. (1982) Clinical and biochemical findings on a child with multiple biotin-responsive carboxylase deficiencies, *J Inherited Metab Dis*, 5, 67-68.

Nishimura, T., Iino, T., Mitsuya, M., Bamba, M., Watanabe, H., Tsukahara, D., Kamata, K., Sasaki, K., Ohyama, S., Hosaka, H., Futamura, M., Nagata, Y. and Eiki J. (2009) Identification of novel and potent 2-amino benzamide derivatives as allosteric glucokinase activators, *Bioorganic & Med Chem Let*, 19, 1357-1360.

Otomo, T., Teruya, K., Uegaki, K., Yamazaki, T. and Kyogoku, Y. (1999) Improved segmental isotope labeling of proteins and application to a larger protein, *J Biomol NMR*, 14,105-114.

Pacheco-Alvarez, D., Solorzano-Vargas, R. S. and Del Rio, A. L. (2002) Biotin in metabolism and its relationship to human disease, *Arch Med Res,* 33**,** 439-47.

Pendini, N. R., Bailey, L.M. Booker, G.W., Wilce, M.C., Walce, J.C. and Polyak, S.W. (2008) Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency, *Biochimica et Biophysica Acta,* 1784, 973-982.

Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. (1997) Attenuated T_2 relaxation bymutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc Natl Acad Sci*, 94, 12366-12371.

Polyak, S. W., Chapman-Smith, A., Brautigan, P. J. and Wallace, J. C. (1999) Biotin protein ligase from Saccharomyces cerevisiae. The N-terminal domain is required for complete activity, *J Biol Chem,* 274**,** 32847-54.

Provencher, S. W. and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism, *Biochemistry*, 20, 33– 37.

Ramachandran, G.N>, Ramakrishnan, C. and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations, *J. Mol. Biol.*, 7, 95–99.

Reche, P., Li, Y. L., Fuller, C., Eichhorn, K. and Perham, R. N. (1998) Selectivity of posttranslational modification in biotinylated proteins: The carboxy carrier protein of the acetyl-CoA carboxylase of Escherichia coli, *Biochem J,* 329 (Pt 3)**,** 589-96.

Reche, P. and Perham, R. N. (1999) Structure and selectivity in post-translational modification: Attaching the biotinyl-lysine and lipoyl-lysine swinging arms in multifunctional enzymes, *EMBO J,* 18**,** 2673-82.

Reinhard, C., Hébert, S.S. and De Strooper, B. (2005) The amyloid-β precursor protein: integrating structure with biological function, *EMBO*, 24, 3996-4006.

Roberts, E.L., Shu, N., Howard, M.J., Broadhurst, R.W., Chapman-Smith, A., Wallace, J.C., Morris, T., Cronan, J.E. and 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, 5045-5053.

Roberts, G.C.K. (1993) “NMR of Macromolecules: A Practical Approach.” New York, NY, Oxford University Press.

Sakamoto, O., Suzuki, Y., Li, X., Aoki, Y., Hiratsuka, M., Holme, E., Kudoh, J., Shimizu, N. and Narisawa, K. (2000) Diagnosis and molecular analysis of an atypical case of holocarboxylase synthetase deficiency, *Eur J Pediatr,* 159**,** 18-22.

Sakamoto, O., Suzuki, Y., Li, X., Aoki, Y., Hiratsuka, M., Suormala, T., Baumgartner, E. R., Gibson, K. M. and Narisawa, K. (1999) Relationship between kinetic properties of mutant enzyme and biochemical and clinical responsiveness to biotin in holocarboxylase synthetase deficiency, *Pediatr Res,* 46**,** 671-6.

Sakamoto, O., Suzuki, Y., Aoki, Y., Li, X., Hiratsuka, M., Yanagihara, K., Inui, K., Okabe, T., Yamaguchi, S., Kudoh, J., Shimizu, N. and Narisawa, K. (1998) Molecular analysis of

new Japanese patients with holocarboxylase synthetase deficiency, *J Inherit Metab Dis,* 21**,** 873-4.

Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C. and Wood, H. G. (1988) Evolutionary conservation among biotin enzymes, *J Biol Chem,* 263, 6461-4.

Soares da Costa, T.P. (2010) Biotinol-5´-AMP forms a slow dissociating complex with *S. aureus* BPL, personal communication.

Spronk, C.A., Nabuurs, S.B., Krieger, E., Vriend, G. and Vuister, G.W. (2004) Validation of protein structures derived by NMR spectroscopy, *Progress NMR Spec*, 45, 315-337.

Sreerama, N and Woody, R.W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism, *Anal Biochem*, 209, 32– 44.

Sreerama, N. and Woody, R. W. (2004) Computation and Analysis of Protein Circular Dichroism Spectra, *Meth Enzymol*, 383, 318-351.

Sun, B., Bird, A., Young, S.P., Kishnani, P.S., Chen, Y.T., Koeberl, D.D. (2007) Enhanced response to enzyme replacement therapy in Pompe disease after the induction of immune tolerance, *Am J Hum Genet*, 81, 1042-1049.

Suormala, T., Fowler, B., Jakobs, C., Duran, M., Lehnert, W., Raab, K., Wick, H. and Baumgartner, E.R. (1998) Late-onset holocarboxylase synthetase-deficiency: pre- and post-natal diagnosis and evaluation of effectiveness of antenatal biotin therapy, *Eur J Pediatr*, 157, 570-575.

Stanley, J.S., Griffin, J.B. and Zempleni, J. (2001) Biotinylation of histones in human cells: Effects of cell proliferation, *Eur J Biochem,* 268, 5424-5429.

Stano, N.M. and Patel, S.S. (2004) T7 lysozyme represses T7 RNA polymerase transcription by destabilizing the open complex during initiation, *J Biol Chem*, 279, 16136-16143.

Streaker, E.D. and Beckett, D. (2003) Coupling of protein assembly and DNA binding: biotin repressor dimerization precedes biotin operator binding, *J Mol Biol*, 325, 937-948.

Sun, M-K. and Alkon, D.L. (2002) Carbonic anhydrase gating of attention: memory therapy and enhancement, *TRENDS Pharm Sci*, 23, 83-89.

Sun, M-K. and Alkon, D.L. (2001) Pharmacological enhancement of synaptic efficacy, spatial learning, and memory through carbonic anhydrase activation in rats, *J. Pharmacol. Exp. Ther*, 297, 961–967.

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

Suzuki, Y., Aoki, Y., Sakamoto, O., Li, X., Miyabayashi, S., Kazuta, Y., Kondo, H. and Narisawa, K. (1996) Enzymatic diagnosis of holocarboxylase synthetase deficiency using apo-carboxyl carrier protein as a substrate, *Clin Chim Acta,* 251**,** 41-52.

Suzuki, Y., Yang, X., Aoki, Y., Kure, S. and Matsubara, Y. (2005) Mutations in the holocarboxylase synthetase gene hlcs, *Hum Mutat,* 26**,** 285-90.

Sweetman, L. and Nyhan, W,L. (1986) Inheritable biotin-treatable disorders and associated phenomena, *Annu Rev Nutr*, 6, 317–343.

Swift, R. (2004) Domain Mapping and functional analysis of human biotin protein ligase (Honours thesis) *School of Molecular and Biomedical Science* University of Adelaide, Adelaide

Temperini, C., Scozzafava, A., Puccettib, L. and Supurana,C.T. (2005) Carbonic anhydrase activators: X-ray crystal structure of the adduct of human isozyme II with L-histidine as a platform for the design of stronger activators, *Bioorganic & Medicinal Chem*, 15, 5136-5141.

Tissot, G., Douce, R. and Alban, C. (1997) Evidence for multiple forms of biotin holocarboxylase synthetase in pea (Pisum sativum) and in Arabidopsis thaliana: Subcellular fractionation studies and isolation of a cDNA clone, *Biochem J,* 323 (Pt 1)**,** 179-88.

Van Stokkum, I.H.M., Spoelder, H.J.W., Bloemendal, M., van Grondelle, R. and Groen, F.C.A. (1990) Estimation of protein secondary structure and error analysis from CD spectra, *Anal Biochem*, 191, 110-118.

Venyaminov, S.Y., Baikalov, I.A., Shen, Z.M., Wu, C.S.C. and Yang, J. T. (1993) Circular dichroic analysis of denatured proteins: Inclusion of denatured protein in the reference set, *Anal Biochem*, 214**,** 17–24.

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

Weaver, L.H., Kwon, K., Beckett, D. and Matthews, B.D. (2001) Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator, *Proc Natl Acad Sci*, 98, 6045-6050.

Wilson, C.J., Myer, M, Darlow, B.A., Stanley, T., Thompson, G., Baumgartner, E.R., Kirby, D.M. and Thorburn, D.R. (2005) Severe Holocarboxylase Synthetase Deficiency with Incomplete Biotin Responsiveness Resulting in Antenatal Insult in Samoan Neonates, *Pedriatrics*, 147, 115-118.

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

Whitmore, L. and Wallace, B.A. (2007) Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases, *Biopol,* 89(5), 392-400.

Wolf, B. (1995) Disorders of biotin metabolism: the metabolic and molecular basis of inherited diseases. *McGraw-Hill*, 7, 3151-3177.

Wolf, B. and Feldman, G.L. (1982) The biotin-dependent carboxylase deficiensies, *Am J Hum Gen*, 34, 699-716.

Wolf, B., Hsia, Y.E., Sweetman, L. and Feldman, G.L. (1981) Multiple carboxylase deficiency: Clinical and biochemical improvement following neonatal biotin treatment, *Pediratrics*, 68, 113-118.

Wuthrich, K. (1986) NMR of proteins and nucleic acids, New York, *Wiley-Interscience*.

Wyatt, P.J. (1993) Light scattering and the absolute characterization of macromolecules, *Anal Chim Acta*, 272, 1-40.

Xu, Y. and Beckett, D. (1994) Kinetics of Biotinyl-5´-adenylate Synthesis Catalyzed by the *Escherichia coli* Repressor of Biotin Biosynthesis and the Stability of the Enzyme-Product Complex, *Biochemistry*, 33, 7354-7360.

Yang, X., Aoki, Y., Li, X., Sakamoto, O., Hiratsuka, M., Kure, S., Taheri, S., Christensen, E., Inui, K., Kubota, M., Ohira, M., Ohki, M., Kudoh, J., Kawasaki, K., Shibuya, K., Shintani,

A., Asakawa, S., Minoshima, S., Shimizu, N., Narisawa, K., Matsubara, Y. and Suzuki, Y. (2001) Structure of human holocarboxylase synthetase gene and mutation spectrum of holocarboxylase synthetase deficiency, *Hum Genet*, 109, 526-534.

Yokoi, K., Ito, T., Maeda, Y., Nakajima, Y., Kurono, Y., Sugiyama, N. and Togari, H. (2009) A case of holocarboxylase synthetase deficiency with insufficient response to prenatal biotin therapy, *Brain & Develop*, 31, 775-778.

Zempleni, J. and Mock, D.M. (2000) Marginal biotin deficiency is teratogenic, *Proc Soc Exp Biol Med*, 223, 14-21.